UNITED STATES AIR FORCE

SUMMER RESEARCH PROGRAM -- 1996

GRADUATE STUDENT RESEARCH PROGRAM FINAL REPORTS

VOLUME 10A

WRIGHT LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES

5800 Uplander Way

Culver City, CA 90230-6608

Program Director, RDL
Gary Moore

Program Manager, AFOSR
Major Linda Steel-Goodwin

Program Manager, RDL
Scott Licoscos

Program Administrator, RDL
Johnetta Thompson

Program Administrator, RDL
Rebecca Kelly

Submitted to:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Bolling Air Force Base

Washington, D.C.

December 1996

REPORT DOCUMENTATION PAGE

AFRL-SR-BL-TR-00-

0740

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruc... ating and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, incl... ate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Bu...

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE December, 1996	3.

4. TITLE AND SUBTITLE 1996 Summer Research Program (SRP), Graduate Student Research Program (GSRP), Final Reports, Volume 10A, Wright Laboratory	5. FUNDING NUMBERS F49620-93-C-0063
6. AUTHOR(S) Gary Moore	

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Research & Development Laboratories (RDL) 5800 Uplander Way Culver City, CA 90230-6608	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Office of Scientific Research (AFOSR) 801 N. Randolph St. Arlington, VA 22203-1977	10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT Approved for Public Release	12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The United States Air Force Summer Research Program (USAF-SRP) is designed to introduce university, college, and technical institute faculty members, graduate students, and high school students to Air Force research. This is accomplished by the faculty members (Summer Faculty Research Program, (SFRP)), graduate students (Graduate Student Research Program (GSRP)), and high school students (High School Apprenticeship Program (HSAP)) being selected on a nationally advertised competitive basis during the summer intersession period to perform research at Air Force Research Laboratory (AFRL) Technical Directorates, Air Force Air Logistics Centers (ALC), and other AF Laboratories. This volume consists of a program overview, program management statistics, and the final technical reports from the GSRP participants at the Wright Laboratory.

14. SUBJECT TERMS Air Force Research, Air Force, Engineering, Laboratories, Reports, Summer, Universities, Faculty, Graduate Student, High School Student	15. NUMBER OF PAGES
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL

Standard Form 298 (Rev. 2-89) (EG)
Prescribed by ANSI Std. 239.18
Designed using Perform Pro, WHS/DIOR, Oct 94

PREFACE

Reports in this volume are numbered consecutively beginning with number 1. Each report is paginated with the report number followed by consecutive page numbers, e.g., 1-1, 1-2, 1-3; 2-1, 2-2, 2-3.

Due to its length, Volume 10 is bound in two parts, 10A, and 10B. Volume 10A contains #1-22. Volume 10B contains reports #23-35.   The Table of Contents for Volume 10 is included in all parts.

This document is one of a set of 16 volumes describing the 1996 AFOSR Summer Research Program. The following volumes comprise the set:

VOLUME	TITLE
1	Program Management Report
	Summer Faculty Research Program (SFRP) Reports
2A & 2B	Armstrong Laboratory
3A & 3B	Phillips Laboratory
4	Rome Laboratory
5A , 5B & 5C	Wright Laboratory
6	Arnold Engineering Development Center, Wilford Hall Medical Center and Air Logistics Centers
	Graduate Student Research Program (GSRP) Reports
7A & 7B	Armstrong Laboratory
8	Phillips Laboratory
9	Rome Laboratory
10A & 10B	Wright Laboratory
11	Arnold Engineering Development Center, United States Air Force Academy, Wilford Hall Medical Center, and Wright Patterson Medical Center
	High School Apprenticeship Program (HSAP) Reports
12A & 12B	Armstrong Laboratory
13	Phillips Laboratory
14	Rome Laboratory
15A&15B	Wright Laboratory
16	Arnold Engineering Development Center

GSRP FINAL REPORT TABLE OF CONTENTS i-x

GSRP FINAL REPORTS

SRP Final Report Table of Contents

SRP Final Report Table of Contents

Author	University/Institution Report Title	Armstrong Laboratory Directorate	Vol-Page
MR Jeremy D Schaub		AL/AOCY	7 - 27
	University of Texas at Austin, Austin, TX		
	In Vitro Evaluation of Lumped Parameter Arterial Models of the Cardiovascular System		
MS Nicole L Schneider		AL/HRGO	7 - 28
	Wright State University, Dayton, OH		
	Java-Based Application of the Model-View-Controller Framework in Developing Interfaces to Interactive Simulations		
MR Christopher S Schreiner		AL/HRA	7 - 29
	Miami University, Oxford, OH		
	The Ability to Reproduce Projective Invariants of Conics		
MS Jacqueline C Shin		AL/HRM	7 - 30
	Pennsylvania State University, University Park, PA		
	Arithmetic Effects on aiming Performance in Coordination: Sequential Position Effects		
MS Emily B Skitek		AL/OER	7 - 31
	Trinity University, San Antonio, TX		
	Does Nitric Oxide Mediate Circulatory Failure Induced by Environmental Heating?		
MR Travis C Tubre		AL/HRT	7 - 32
	Texas A&M University College station, College Station, TX		
	The Development of A General Measure of Performance		
MR Reynardo D Tyner		AL/CFBV	7 - 33
	Auburn University, Auburn, AL		
MR Christopher G Walker		AL/EQC	7 - 34
	Jackson State University, Jackson, MS		
	The Analysis of Aqueous Film Forming Foam		
MR Ross E Willis		AL/HRTI	7 - 35
	Texas Tech University, Lubbock, TX		
	Automating the Cognitive Task Analysis Procedure		

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

INTRODUCTION

The Summer Research Program (SRP), sponsored by the Air Force Office of Scientific Research (AFOSR), offers paid opportunities for university faculty, graduate students, and high school students to conduct research in U.S. Air Force research laboratories nationwide during the summer.

Introduced by AFOSR in 1978, this innovative program is based on the concept of teaming academic researchers with Air Force scientists in the same disciplines using laboratory facilities and equipment not often available at associates' institutions.

The Summer Faculty Research Program (SFRP) is open annually to approximately 150 faculty members with at least two years of teaching and/or research experience in accredited U.S. colleges, universities, or technical institutions. SFRP associates must be either U.S. citizens or permanent residents.

The Graduate Student Research Program (GSRP) is open annually to approximately 100 graduate students holding a bachelor's or a master's degree; GSRP associates must be U.S. citizens enrolled full time at an accredited institution.

The High School Apprentice Program (HSAP) annually selects about 125 high school students located within a twenty mile commuting distance of participating Air Force laboratories.

AFOSR also offers its research associates an opportunity, under the Summer Research Extension Program (SREP), to continue their AFOSR-sponsored research at their home institutions through the award of research grants. In 1994 the maximum amount of each grant was increased from $20,000 to $25,000, and the number of AFOSR-sponsored grants decreased from 75 to 60. A separate annual report is compiled on the SREP.

The numbers of projected summer research participants in each of the three categories and SREP "grants" are usually increased through direct sponsorship by participating laboratories.

AFOSR's SRP has well served its objectives of building critical links between Air Force research laboratories and the academic community, opening avenues of communications and forging new research relationships between Air Force and academic technical experts in areas of national interest, and strengthening the nation's efforts to sustain careers in science and engineering. The success of the SRP can be gauged from its growth from inception (see Table 1) and from the favorable responses the 1996 participants expressed in end-of-tour SRP evaluations (Appendix B).

AFOSR contracts for administration of the SRP by civilian contractors. The contract was first awarded to Research & Development Laboratories (RDL) in September 1990. After

1

completion of the 1990 contract, RDL (in 1993) won the recompetition for the basic year and four 1-year options.

2. PARTICIPATION IN THE SUMMER RESEARCH PROGRAM

The SRP began with faculty associates in 1979; graduate students were added in 1982 and high school students in 1986. The following table shows the number of associates in the program each year.

YEAR	SRP Participation, by Year			TOTAL
	SFRP	GSRP	HSAP	
1979	70			70
1980	87			87
1981	87			87
1982	91	17		108
1983	101	53		154
1984	152	84		236
1985	154	92		246
1986	158	100	42	300
1987	159	101	73	333
1988	153	107	101	361
1989	168	102	103	373
1990	165	121	132	418
1991	170	142	132	444
1992	185	121	159	464
1993	187	117	136	440
1994	192	117	133	442
1995	190	115	137	442
1996	188	109	138	435

Beginning in 1993, due to budget cuts, some of the laboratories weren't able to afford to fund as many associates as in previous years. Since then, the number of funded positions has remained fairly constant at a slightly lower level.

3. RECRUITING AND SELECTION

The SRP is conducted on a nationally advertised and competitive-selection basis. The advertising for faculty and graduate students consisted primarily of the mailing of 8,000 52-page SRP brochures to chairpersons of departments relevant to AFOSR research and to administrators of grants in accredited universities, colleges, and technical institutions. Historically Black Colleges and Universities (HBCUs) and Minority Institutions (MIs) were included. Brochures also went to all participating USAF laboratories, the previous year's participants, and numerous individual requesters (over 1000 annually).

RDL placed advertisements in the following publications: *Black Issues in Higher Education, Winds of Change,* and *IEEE Spectrum.* Because no participants list either *Physics Today* or *Chemical & Engineering News* as being their source of learning about the program for the past several years, advertisements in these magazines were dropped, and the funds were used to cover increases in brochure printing costs.

High school applicants can participate only in laboratories located no more than 20 miles from their residence. Tailored brochures on the HSAP were sent to the head counselors of 180 high schools in the vicinity of participating laboratories, with instructions for publicizing the program in their schools. High school students selected to serve at Wright Laboratory's Armament Directorate (Eglin Air Force Base, Florida) serve eleven weeks as opposed to the eight weeks normally worked by high school students at all other participating laboratories.

Each SFRP or GSRP applicant is given a first, second, and third choice of laboratory. High school students who have more than one laboratory or directorate near their homes are also given first, second, and third choices.

Laboratories make their selections and prioritize their nominees. AFOSR then determines the number to be funded at each laboratory and approves laboratories' selections.

Subsequently, laboratories use their own funds to sponsor additional candidates. Some selectees do not accept the appointment, so alternate candidates are chosen. This multi-step selection procedure results in some candidates being notified of their acceptance after scheduled deadlines. The total applicants and participants for 1996 are shown in this table.

1996 Applicants and Participants			
PARTICIPANT CATEGORY	TOTAL APPLICANTS	SELECTEES	DECLINING SELECTEES
SFRP	572	188	39
(HBCU/MI)	(119)	(27)	(5)
GSRP	235	109	7
(HBCU/MI)	(18)	(7)	(1)
HSAP	474	138	8
TOTAL	1281	435	54

4. SITE VISITS

During June and July of 1996, representatives of both AFOSR/NI and RDL visited each participating laboratory to provide briefings, answer questions, and resolve problems for both laboratory personnel and participants. The objective was to ensure that the SRP would be as constructive as possible for all participants. Both SRP participants and RDL representatives found these visits beneficial. At many of the laboratories, this was the only opportunity for all participants to meet at one time to share their experiences and exchange ideas.

5. HISTORICALLY BLACK COLLEGES AND UNIVERSITIES AND MINORITY INSTITUTIONS (HBCU/MIs)

Before 1993, an RDL program representative visited from seven to ten different HBCU/Mis annually to promote interest in the SRP among the faculty and graduate students. These efforts were marginally effective, yielding a doubling of HBCI/MI applicants. In an effort to achieve AFOSR's goal of 10% of all applicants and selectees being HBCU/MI qualified, the RDL team decided to try other avenues of approach to increase the number of qualified applicants. Through the combined efforts of the AFOSR Program Office at Bolling AFB and RDL, two very active minority groups were found, HACU (Hispanic American Colleges and Universities) and AISES (American Indian Science and Engineering Society). RDL is in communication with representatives of each of these organizations on a monthly basis to keep up with the their activities and special events. Both organizations have widely-distributed magazines/quarterlies in which RDL placed ads.

Since 1994 the number of both SFRP and GSRP HBCU/MI applicants and participants has increased ten-fold, from about two dozen SFRP applicants and a half dozen selectees to over 100 applicants and two dozen selectees, and a half-dozen GSRP applicants and two or three selectees to 18 applicants and 7 or 8 selectees. Since 1993, the SFRP had a two-fold applicant

4

increase and a two-fold selectee increase. Since 1993, the GSRP had a three-fold applicant increase and a three to four-fold increase in selectees.

In addition to RDL's special recruiting efforts, AFOSR attempts each year to obtain additional funding or use leftover funding from cancellations the past year to fund HBCU/MI associates. This year, 5 HBCU/MI SFRPs declined after they were selected (and there was no one qualified to replace them with). The following table records HBCU/MI participation in this program.

SRP HBCU/MI Participation, By Year				
YEAR	SFRP		GSRP	
	Applicants	Participants	Applicants	Participants
1985	76	23	15	11
1986	70	18	20	10
1987	82	32	32	10
1988	53	17	23	14
1989	39	15	13	4
1990	43	14	17	3
1991	42	13	8	5
1992	70	13	9	5
1993	60	13	6	2
1994	90	16	11	6
1995	90	21	20	8
1996	119	27	18	7

6. SRP FUNDING SOURCES

Funding sources for the 1996 SRP were the AFOSR-provided slots for the basic contract and laboratory funds. Funding sources by category for the 1996 SRP selected participants are shown here.

1996 SRP FUNDING CATEGORY	SFRP	GSRP	HSAP
AFOSR Basic Allocation Funds	141	85	123
USAF Laboratory Funds	37	19	15
HBCU/MI By AFOSR (Using Procured Addn'l Funds)	10	5	0
TOTAL	188	109	138

SFRP - 150 were selected, but nine canceled too late to be replaced.
GSRP - 90 were selected, but five canceled too late to be replaced (10 allocations for the ALCs were withheld by AFOSR.)
HSAP - 125 were selected, but two canceled too late to be replaced.

7. COMPENSATION FOR PARTICIPANTS

Compensation for SRP participants, per five-day work week, is shown in this table.

1996 SRP Associate Compensation

PARTICIPANT CATEGORY	1991	1992	1993	1994	1995	1996
Faculty Members	$690	$718	$740	$740	$740	$770
Graduate Student (Master's Degree)	$425	$442	$455	$455	$455	$470
Graduate Student (Bachelor's Degree)	$365	$380	$391	$391	$391	$400
High School Student (First Year)	$200	$200	$200	$200	$200	$200
High School Student (Subsequent Years)	$240	$240	$240	$240	$240	$240

The program also offered associates whose homes were more than 50 miles from the laboratory an expense allowance (seven days per week) of $50/day for faculty and $40/day for graduate students. Transportation to the laboratory at the beginning of their tour and back to their home destinations at the end was also reimbursed for these participants. Of the combined SFRP and

6

GSRP associates, 65 % (194 out of 297) claimed travel reimbursements at an average round-trip cost of $780.

Faculty members were encouraged to visit their laboratories before their summer tour began. All costs of these orientation visits were reimbursed. Forty-five percent (85 out of 188) of faculty associates took orientation trips at an average cost of $444. By contrast, in 1993, 58 % of SFRP associates took orientation visits at an average cost of $685; that was the highest percentage of associates opting to take an orientation trip since RDL has administered the SRP, and the highest average cost of an orientation trip. These 1993 numbers are included to show the fluctuation which can occur in these numbers for planning purposes.

Program participants submitted biweekly vouchers countersigned by their laboratory research focal point, and RDL issued paychecks so as to arrive in associates' hands two weeks later.

In 1996, RDL implemented direct deposit as a payment option for SFRP and GSRP associates. There were some growing pains. Of the 128 associates who opted for direct deposit, 17 did not check to ensure that their financial institutions could support direct deposit (and they couldn't), and eight associates never did provide RDL with their banks' ABA number (direct deposit bank routing number), so only 103 associates actually participated in the direct deposit program. The remaining associates received their stipend and expense payments via checks sent in the US mail.

HSAP program participants were considered actual RDL employees, and their respective state and federal income tax and Social Security were withheld from their paychecks. By the nature of their independent research, SFRP and GSRP program participants were considered to be consultants or independent contractors. As such, SFRP and GSRP associates were responsible for their own income taxes, Social Security, and insurance.

8. CONTENTS OF THE 1996 REPORT

The complete set of reports for the 1996 SRP includes this program management report (Volume 1) augmented by fifteen volumes of final research reports by the 1996 associates, as indicated below:

1996 SRP Final Report Volume Assignments

LABORATORY	SFRP	GSRP	HSAP
Armstrong	2	7	12
Phillips	3	8	13
Rome	4	9	14
Wright	5A, 5B	10	15
AEDC, ALCs, WHMC	6	11	16

APPENDIX A – PROGRAM STATISTICAL SUMMARY

A. Colleges/Universities Represented

Selected SFRP associates represented 169 different colleges, universities, and institutions, GSRP associates represented 95 different colleges, universities, and institutions.

B. States Represented

SFRP -Applicants came from 47 states plus Washington D.C. and Puerto Rico. Selectees represent 44 states plus Puerto Rico.

GSRP - Applicants came from 44 states and Puerto Rico. Selectees represent 32 states.

HSAP - Applicants came from thirteen states. Selectees represent nine states.

Total Number of Participants	
SFRP	188
GSRP	109
HSAP	138
TOTAL	435

Degrees Represented			
	SFRP	GSRP	TOTAL
Doctoral	184	1	185
Master's	4	48	52
Bachelor's	0	60	60
TOTAL	188	109	297

SFRP Academic Titles	
Assistant Professor	79
Associate Professor	59
Professor	42
Instructor	3
Chairman	0
Visiting Professor	1
Visiting Assoc. Prof.	0
Research Associate	4
TOTAL	188

Source of Learning About the SRP		
Category	Applicants	Selectees
Applied/participated in prior years	28%	34%
Colleague familiar with SRP	19%	16%
Brochure mailed to institution	23%	17%
Contact with Air Force laboratory	17%	23%
IEEE Spectrum	2%	1%
BIIHE	1%	1%
Other source	10%	8%
TOTAL	100%	100%

APPENDIX B – SRP EVALUATION RESPONSES

1. OVERVIEW

Evaluations were completed and returned to RDL by four groups at the completion of the SRP. The number of respondents in each group is shown below.

Table B-1. Total SRP Evaluations Received

Evaluation Group	Responses
SFRP & GSRPs	275
HSAPs	113
USAF Laboratory Focal Points	84
USAF Laboratory HSAP Mentors	6

All groups indicate unanimous enthusiasm for the SRP experience.

The summarized recommendations for program improvement from both associates and laboratory personnel are listed below:

A.   Better preparation on the labs' part prior to associates' arrival (i.e., office space, computer assets, clearly defined scope of work).

B.   Faculty Associates suggest higher stipends for SFRP associates.

C.   Both HSAP Air Force laboratory mentors and associates would like the summer tour extended from the current 8 weeks to either 10 or 11 weeks; the groups state it takes 4-6 weeks just to get high school students up-to-speed on what's going on at laboratory. (Note: this same argument was used to raise the faculty and graduate student participation time a few years ago.)

2. 1996 USAF LABORATORY FOCAL POINT (LFP) EVALUATION RESPONSES

The summarized results listed below are from the 84 LFP evaluations received.

1. LFP evaluations received and associate preferences:

Table B-2. Air Force LFP Evaluation Responses (By Type)

| | | How Many Associates Would You Prefer To Get ? (% Response) | | | | | | | | | | | |
| | | SFRP | | | | GSRP (w/Univ Professor) | | | | GSRP (w/o Univ Professor) | | | |
Lab	Evals Recv'd	0	1	2	3+	0	1	2	3+	0	1	2	3+
AEDC	0	-	-	-	-	-	-	-	-	-	-	-	-
WHMC	0	-	-	-	-	-	-	-	-	-	-	-	-
AL	7	28	28	28	14	54	14	28	0	86	0	14	0
FJSRL	1	0	100	0	0	100	0	0	0	0	100	0	0
PL	25	40	40	16	4	88	12	0	0	84	12	4	0
RL	5	60	40	0	0	80	10	0	0	100	0	0	0
WL	46	30	43	20	6	78	17	4	0	93	4	2	0
Total	84	32%	50%	13%	5%	80%	11%	6%	0%	73%	23%	4%	0%

LFP Evaluation Summary. The summarized responses, by laboratory, are listed on the following page. LFPs were asked to rate the following questions on a scale from 1 (below average) to 5 (above average).

2. LFPs involved in SRP associate application evaluation process:
    a. Time available for evaluation of applications:
    b. Adequacy of applications for selection process:
3. Value of orientation trips:
4. Length of research tour:
5   a. Benefits of associate's work to laboratory:
    b. Benefits of associate's work to Air Force:
6.   a. Enhancement of research qualifications for LFP and staff:
    b. Enhancement of research qualifications for SFRP associate:
    c. Enhancement of research qualifications for GSRP associate:
7.   a. Enhancement of knowledge for LFP and staff:
    b. Enhancement of knowledge for SFRP associate:
    c. Enhancement of knowledge for GSRP associate:
8. Value of Air Force and university links:
9. Potential for future collaboration:
10.   a. Your working relationship with SFRP:
    b. Your working relationship with GSRP:
11. Expenditure of your time worthwhile:

12. Quality of program literature for associate:
13.   a. Quality of RDL's communications with you:
        b. Quality of RDL's communications with associates:
14. Overall assessment of SRP:

Table B-3. Laboratory Focal Point Reponses to above questions

	AEDC	AL	FJSRL	PL	RL	WHMC	WL
# Evals Recv'd	0	7	1	14	5	0	46
Question #							
2	-	86 %	0 %	88 %	80 %	-	85 %
2a	-	4.3	n/a	3.8	4.0	-	3.6
2b	-	4.0	n/a	3.9	4.5	-	4.1
3	-	4.5	n/a	4.3	4.3	-	3.7
4	-	4.1	4.0	4.1	4.2	-	3.9
5a	-	4.3	5.0	4.3	4.6	-	4.4
5b	-	4.5	n/a	4.2	4.6	-	4.3
6a	-	4.5	5.0	4.0	4.4	-	4.3
6b	-	4.3	n/a	4.1	5.0	-	4.4
6c	-	3.7	5.0	3.5	5.0	-	4.3
7a	-	4.7	5.0	4.0	4.4	-	4.3
7b	-	4.3	n/a	4.2	5.0	-	4.4
7c	-	4.0	5.0	3.9	5.0	-	4.3
8	-	4.6	4.0	4.5	4.6	-	4.3
9	-	4.9	5.0	4.4	4.8	-	4.2
10a	-	5.0	n/a	4.6	4.6	-	4.6
10b	-	4.7	5.0	3.9	5.0	-	4.4
11	-	4.6	5.0	4.4	4.8	-	4.4
12	-	4.0	4.0	4.0	4.2	-	3.8
13a	-	3.2	4.0	3.5	3.8	-	3.4
13b	-	3.4	4.0	3.6	4.5	-	3.6
14	-	4.4	5.0	4.4	4.8	-	4.4

3. 1996 SFRP & GSRP EVALUATION RESPONSES

The summarized results listed below are from the 257 SFRP/GSRP evaluations received.

Associates were asked to rate the following questions on a scale from 1 (below average) to 5 (above average) - by Air Force base results and over-all results of the 1996 evaluations are listed after the questions.

1. The match between the laboratories research and your field:
2. Your working relationship with your LFP:
3. Enhancement of your academic qualifications:
4. Enhancement of your research qualifications:
5. Lab readiness for you: LFP, task, plan:
6. Lab readiness for you: equipment, supplies, facilities:
7. Lab resources:
8. Lab research and administrative support:
9. Adequacy of brochure and associate handbook:
10. RDL communications with you:
11. Overall payment procedures:
12. Overall assessment of the SRP:
13.   a. Would you apply again?
       b. Will you continue this or related research?
14. Was length of your tour satisfactory?
15. Percentage of associates who experienced difficulties in finding housing:
16. Where did you stay during your SRP tour?
       a.   At Home:
       b.   With Friend:
       c.   On Local Economy:
       d.   Base Quarters:
17. Value of orientation visit:
       a.   Essential:
       b.   Convenient:
       c.   Not Worth Cost:
       d.   Not Used:

SFRP and GSRP associate's responses are listed in tabular format on the following page.

Table B-4. 1996 SFRP & GSRP Associate Responses to SRP Evaluation

	Arnold	Brooks	Edwards	Eglin	Griffis	Hanscom	Kelly	Kirtland	Lackland	Robins	Tyndall	WPAFB	average
# res	6	48	6	14	31	19	3	32	1	2	10	85	257
1	4.8	4.4	4.6	4.7	4.4	4.9	4.6	4.6	5.0	5.0	4.0	4.7	4.6
2	5.0	4.6	4.1	4.9	4.7	4.7	5.0	4.7	5.0	5.0	4.6	4.8	4.7
3	4.5	4.4	4.0	4.6	4.3	4.2	4.3	4.4	5.0	5.0	4.5	4.3	4.4
4	4.3	4.5	3.8	4.6	4.4	4.4	4.3	4.6	5.0	4.0	4.4	4.5	4.5
5	4.5	4.3	3.3	4.8	4.4	4.5	4.3	4.2	5.0	5.0	3.9	4.4	4.4
6	4.3	4.3	3.7	4.7	4.4	4.5	4.0	3.8	5.0	5.0	3.8	4.2	4.2
7	4.5	4.4	4.2	4.8	4.5	4.3	4.3	4.1	5.0	5.0	4.3	4.3	4.4
8	4.5	4.6	3.0	4.9	4.4	4.3	4.3	4.5	5.0	5.0	4.7	4.5	4.5
9	4.7	4.5	4.7	4.5	4.3	4.5	4.7	4.3	5.0	5.0	4.1	4.5	4.5
10	4.2	4.4	4.7	4.4	4.1	4.1	4.0	4.2	5.0	4.5	3.6	4.4	4.3
11	3.8	4.1	4.5	4.0	3.9	4.1	4.0	4.0	3.0	4.0	3.7	4.0	4.0
12	5.7	4.7	4.3	4.9	4.5	4.9	4.7	4.6	5.0	4.5	4.6	4.5	4.6
Numbers below are percentages													
13a	83	90	83	93	87	75	100	81	100	100	100	86	87
13b	100	89	83	100	94	98	100	94	100	100	100	94	93
14	83	96	100	90	87	80	100	92	100	100	70	84	88
15	17	6	0	33	20	76	33	25	0	100	20	8	39
16a	-	26	17	9	38	23	33	4	-	-	-	30	
16b	100	33	-	40	-	8	-	-	-	-	36	2	
16c	-	41	83	40	62	69	67	96	100	100	64	68	
16d	-	-	-	-	-	-	-	-	-	-	-	0	
17a	-	33	100	17	50	14	67	39	-	50	40	31	35
17b	-	21	-	17	10	14	-	24	-	50	20	16	16
17c	-	-	-	-	10	7	-	-	-	-	-	2	3
17d	100	46	-	66	30	69	33	37	100	-	40	51	46

4. 1996 USAF LABORATORY HSAP MENTOR EVALUATION RESPONSES

Not enough evaluations received (5 total) from Mentors to do useful summary.

5. 1996 HSAP EVALUATION RESPONSES

The summarized results listed below are from the 113 HSAP evaluations received.

HSAP apprentices were asked to rate the following questions on a scale from
1 (below average) to 5 (above average)

1. Your influence on selection of topic/type of work.
2. Working relationship with mentor, other lab scientists.
3. Enhancement of your academic qualifications.
4. Technically challenging work.
5. Lab readiness for you: mentor, task, work plan, equipment.
6. Influence on your career.
7. Increased interest in math/science.
8. Lab research & administrative support.
9. Adequacy of RDL's Apprentice Handbook and administrative materials.
10. Responsiveness of RDL communications.
11. Overall payment procedures.
12. Overall assessment of SRP value to you.
13. Would you apply again next year?   Yes (92 %)
14. Will you pursue future studies related to this research?   Yes (68 %)
15. Was Tour length satisfactory?   Yes (82 %)

	Arnold	Brooks	Edwards	Eglin	Griffiss	Hanscom	Kirtland	Tyndall	WPAFB	Totals
# resp	5	19	7	15	13	2	7	5	40	113
1	2.8	3.3	3.4	3.5	3.4	4.0	3.2	3.6	3.6	3.4
2	4.4	4.6	4.5	4.8	4.6	4.0	4.4	4.0	4.6	4.6
3	4.0	4.2	4.1	4.3	4.5	5.0	4.3	4.6	4.4	4.4
4	3.6	3.9	4.0	4.5	4.2	5.0	4.6	3.8	4.3	4.2
5	4.4	4.1	3.7	4.5	4.1	3.0	3.9	3.6	3.9	4.0
6	3.2	3.6	3.6	4.1	3.8	5.0	3.3	3.8	3.6	3.7
7	2.8	4.1	4.0	3.9	3.9	5.0	3.6	4.0	4.0	3.9
8	3.8	4.1	4.0	4.3	4.0	4.0	4.3	3.8	4.3	4.2
9	4.4	3.6	4.1	4.1	3.5	4.0	3.9	4.0	3.7	3.8
10	4.0	3.8	4.1	3.7	4.1	4.0	3.9	2.4	3.8	3.8
11	4.2	4.2	3.7	3.9	3.8	3.0	3.7	2.6	3.7	3.8
12	4.0	4.5	4.9	4.6	4.6	5.0	4.6	4.2	4.3	4.5
Numbers below are percentages										
13	60%	95%	100%	100%	85%	100%	100%	100%	90%	92%
14	20%	80%	71%	80%	54%	100%	71%	80%	65%	68%
15	100%	70%	71%	100%	100%	50%	86%	60%	80%	82%

DEVELOPMENT OF A GLOBAL NAVIGATION SATELLITE SYSTEM
SOFTWARE RADIO

Dennis M. Akos
Ph.D. Candidate
Department of Electrical Engineering and Computer Science

Ohio University
Avionics Engineering Center
Athens, OH 45701

DEVELOPMENT OF A GLOBAL NAVIGATION SATELLITE SYSTEM SOFTWARE RADIO

Dennis M. Akos
Ph.D. Candidate
Department of Electrical Engineering and Computer Science
Ohio University

Abstract

A preliminary Global Navigation Satellite System (GNSS) software radio was developed. A software radio has many advantages over the architecture of a traditional receiver. These include a tighter integration between simulation and implementation, a tremendous level of versatility in the final design, and the ability for a single receiver to function as multiple receivers. The key elements for the development of a software radio are analog-to-digital converter (ADC) technology and programmable processing power. Both of these issues have been investigated. The focus of this implementation, a GNSS receiver, is a navigation receiver and will bring all the benefits of the software radio to the navigation community. The preliminary work accomplished in the development of the GNSS software radio is the implementation of the receiver front end, data collection hardware, and data processing algorithms.

DEVELOPMENT OF A GLOBAL NAVIGATION SATELLITE SYSTEM
SOFTWARE RADIO

Dennis M. Akos

Introduction

The software radio describes a receiver in which the majority of the signal processing is accomplished via a programmable microprocessor as opposed to analog or hardwired discrete components. This allows for a tighter integration of simulation and implementation as well as tremendous flexibility in the final design.

The software radio concept is being applied in the design of a GNSS receiver. However, this concept is not limited to the GNSS signal and could be expanded to included other navigation signals in the same radio design. This initial work will bring the benefits of such an implementation to the navigation community.

The paper begins by describing the ideal software radio and details its multiple benefits. The target implementation and the development testbed is characterized along with the necessary design steps. Finally, an informal discussion of various GNSS acquisition methodologies is presented. Described here are those techniques implemented for use with the software radio and validated using actual GNSS data.

Software Radio

There are two primary design goals in developing a software radio. First, the analog-to-digital converter (ADC) should be positioned as close to the antenna as possible in the front end of the receiver. Second, the resulting samples should be processed using a programmable microprocessor. These two principles provide all the benefits associated with the software radio.

Moving the ADC closer to the antenna in the RF front end chain eliminates additional components used in frequency translation. These components include: local oscillators (LO), mixers, and filters, all of which can contribute potential nonlinear effects as well as temperature and age based performance variations. Ideally, the receiver front end would consist of the antenna, amplifier, bandpass filter, and ADC. Frequency translation, since it is impractical to process the signal at RF, is accomplished via bandpass sampling [1].

Bandpass sampling is the process of sampling an information signal based on its bandwidth as opposed to its RF carrier. This has been proposed and implemented successfully with the GPS-SPS transmission [2]. Reference 2 details a front end design consisting of an antenna, amplifiers, filters and an ADC that sampled the 1575.42 MHz RF carrier directly at a rate of 5 MHz and achieved the desired frequency translation via bandpass sampling.

Processing the resulting ADC samples strictly in software provides additional benefits for the software radio concept. First, since all signal processing is accomplished in software there is tighter integration between simulation and actual receiver operation. If the signal degradations can be adequately simulated, performance can

be accurately predicted using the actual discrete signal processing that will occur in the receiver. This is especially true now that the front end contains fewer possible error sources. Second, there is a tremendous level of flexibility in the receiver design since all signal processing is software based. In order to incorporate the latest theoretical developments, costly hardware prototypes no longer need to be fabricated, rather they can be incorporated into the programming and evaluated. Various receiver architectures can be assessed simply by downloading the appropriate software to the target processor.

The front end and the flexibility of achieving all signal processing in software allow the design to serve as multiple radios [3]. Currently, there exist receivers, where the LO is adjusted to downconvert and process different frequency transmissions. These designs, however, are limited in their signal processing as a result of their hardwired architectures. With the software radio design a continuous range of frequencies could be captured by using a high sampling rate. Specific transmissions could then be digitally filtered out and processed. By changing the software processing, a single configuration could serve as an FM, AM, or PM receiver. This concept would be extremely beneficial in the navigation community as a single receiver could process and integrate multiple navigation signals for improved accuracy, reliability, and integrity.

Some of these ideas are reflected in the current software radio research. Reference 4 discusses a GPS L1/L2 front end design which bandpass samples the frequency band 1.2 - 1.6 GHz to utilize both GPS frequency transmissions. This front end is followed by digital filters used to extract the exact information bands of interest for further processing.

The software radio concept is not without its disadvantages, unfortunately. There are two primary technological factors which limit its current practicality. They are the current state-of-the-art in ADC technology and programmable processing power. For the bulk of the navigation community, the highest RF signals of interest are for the GNSS band which are below 1.7 GHz. ADC's do exist which can provide multi-bit sampling at rates up to 4 Gsps [5]. At this sampling rate all frequency information from DC to 2 GHz can be captured. However, programmable processing power significantly reduces the maximum possible data rate from that allowed by the ADC [6]. This leads to the capture of partial frequency bands, as in the case of bandpass sampling. The limits imposed by the current generation of programmable processors are so restrictive, a variation of bandpass sampling has been proposed for the combined digitization and processing of GPS-SPS and GLONASS signals [7].

These disadvantages are only temporary. ADC performance already exceeds what is required for the navigation community. The lag in available processor power should be eliminated in the near future. Moore's law, which has held true since the inception of the microprocessor, has shown processing power to double every 18 months.

GNSS SOFTWARE RADIO DEVELOPMENT

A navigation software radio will provide significant advantages over existing receivers and their traditional design. Although the current technology will not allow the development of an all-encompassing

1-4
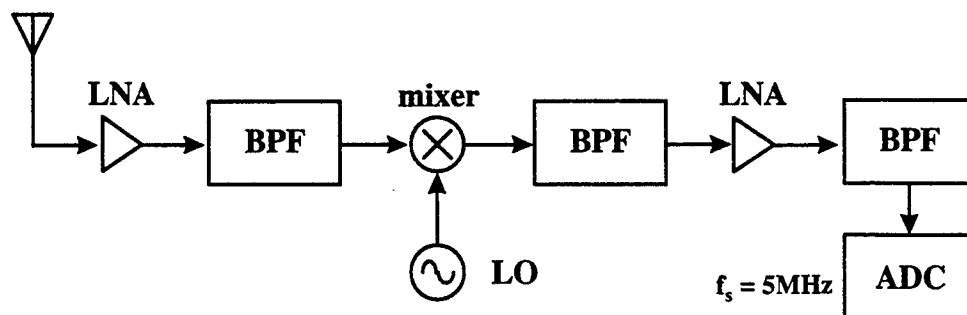
Figure 1. Front End Configuration used for Raw Data Collection

navigation software radio, the initial goal is the development of a GNSS software radio. As technology advances, the framework (front end hardware and software algorithms) will be in place to take advantage of increased processing power.

The development is planned for three stages. First, a front end utilizing multiband bandpass sampling was designed and implemented [7]. This research demonstrated proof of concept and a final design is under development. The second stage, currently underway, is the programming of the software algorithms necessary for processing of the sampled data. It is impractical to attempt to initially develop the algorithms to operate in real time. Rather a data set will be collected and postprocessed using the developing algorithms. The third stage will be the optimization of these algorithms to operate in real time. The target processor for the real time implementation is the Texas Instruments TMS320C80 DSP, one of the most powerful DPS processors available, capable of 2 BOPS.

In order to validate the spread spectrum acquisition and tracking algorithms it is necessary to obtain an adequate length data set. The difficulty lies in the pre-correlation bandwidth of the GNSS signal. GPS-SPS has a null-to-null bandwidth of approximately 2 MHz, therefore the minimum sampling frequency must be at least 4 MHz. A sampling frequency of 5 MHz is used to adequately capture the required frequency information. Assuming 8 bit samples, 30 seconds of data (a full navigation frame for GPS-SPS) requires 150 MB of storage space. In order to minimize the storage requirement, data sets of 12 seconds (or a two subframes) are collected for postprocessing. If subframes #1, 2, and 3 can be collected, the resulting data will contain enough information to establish a position solution, which is the primary purpose in a navigation receiver. This can be accomplish by collecting a single data set corresponding to the desired subframe and then storing that on a hard drive or an alternative long term storage device, then collecting the next desired subframe soon after. This process is repeated until all required subframes can be collected.

The data collection platform uses a more traditional front end design to reduce the requirement on the ADC. The configuration, depicted in Figure 1, employs a single downconversion stage to 21.25 MHz, where the
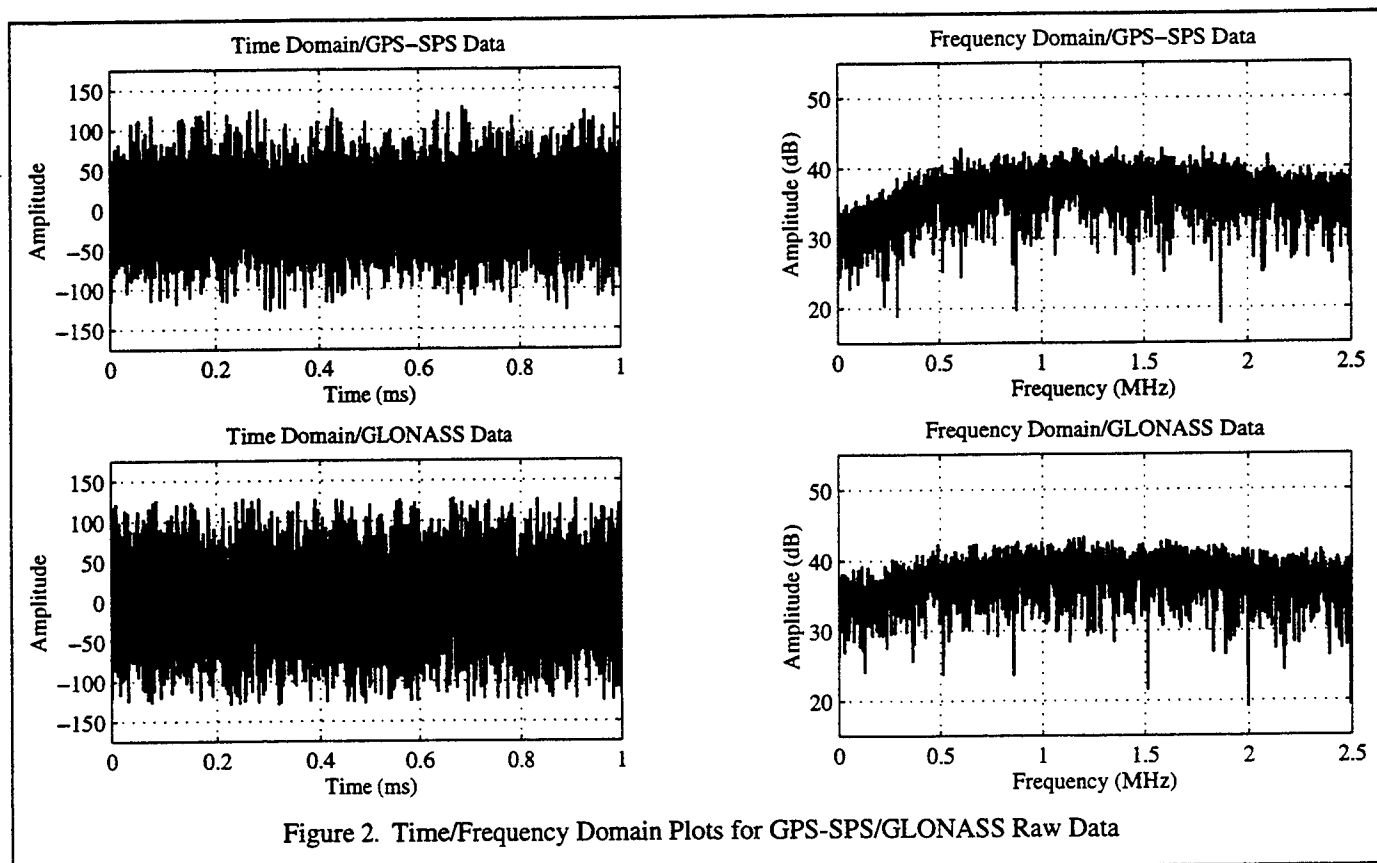
Figure 2. Time/Frequency Domain Plots for GPS-SPS/GLONASS Raw Data

signal is bandpass sampled at 5 MHz, resulting in a final IF of 1.25 MHz. This arrangement is utilized to collect GPS-SPS data sets and the development of generic GNSS signal processing algorithms.

GNSS SOFTWARE RADIO ACQUISITION

The first stage in processing the code division multiple access (CDMA) format of the GPS-SPS navigation signal is acquisition. The spread spectrum modulation format essentially conceals any discernible signal in the raw data set when viewed in either the time or frequency domains. This is also true for GLONASS even though it employs frequency division multiple access (FDMA). Each frequency channel uses the same maximal length code as a spreading sequence for the purpose of time transfer. Raw data collected using the front end in Figure 1 is shown in Figure 2 in both the time and frequency domain. The rolloff in the frequency domain plots is a result of the 2.0 MHz 3 dB bandwidth of the final filter in the RF chain. The raw GPS-SPS data contains the CDMA broadcasts of 5 visible satellites. This same front end configuration allowed enough bandwidth for data capture of 2 of the GLONASS channels. GLONASS data was collected by adjusting the LO of the front end to translate channels 21 & 22 to the resulting sampled bandwidth.

Acquisition is the search for the parameters necessary to identify the signal and begin tracking. In the case of GPS-SPS, this includes the signal's spreading (Coarse/Acquisition (C/A)) code, carrier frequency, and code

phase. GLONASS reduces the search space by one parameter as it uses the same spreading sequence on each frequency channel. The search can be visualized as a matrix (2-D for GLONASS and 3-D for GPS) where every entry must be tested until one is found corresponding to the correct set of parameters. This search space must be bounded with a defined step size. For GPS-SPS there are 32 possible C/A codes. The possible carrier frequency, which differs as a result of Doppler, is bounded for most users to ±10 kHz from nominal and is searched in 500 Hz bins. Lastly the spreading code is 1023 chips for GPS-SPS and 511 chips for GLONASS and is searched in ½ and ¼ chip increments, respectively, over a single code period.

There are a number of popular spread spectrum signal acquisition algorithms [8]. However, most commercial GPS-SPS receivers tend to use the serial search technique. The popularity of this technique is most likely due to the fact that the digital correlator/accumulator hardware can be used not only for tracking, but also acquisition if serial search is employed. In order to demonstrate the flexibility of the software radio approach, multiple acquisition algorithms have been coded.

The postprocessing approach allowed for a more comprehensive evaluation of each of the acquisition techniques. First, the code phase search was stepped in terms of samples. Second, the search conducted was exhaustive, that is every point in the search space was evaluated. In a traditional receiver, points in the search space are sequentially tested until a threshold is crossed indicating a potential match has been obtained and control is transferred to attempt tracking. In serial search there are well-defined equations to calculate the threshold that also determines the associated probabilities (missed detection and false acquisition) [9].

In the standard serial search routine, the signal is converted to baseband using a frequency entry from the test matrix and multiplied by the spreading code with a code phase entry from the test matrix. The resulting data points are accumulated over a single code period and that measurement is used to determine if the correct entry from the matrix has been found. Although this is a well-established technique, the disadvantage is that all test points in the matrix must be evaluated serially, as implied by the name. From the earlier discussion exhaustive testing of single C/A code or GLONASS frequency will require evaluating:

$$\left(\frac{10000}{500}\right)\left(\frac{1023}{\frac{1}{2}}\right) \approx \left(\frac{10000}{500}\right)\left(\frac{511}{\frac{1}{4}}\right) \approx 40920 \qquad (1)$$

possible entries in the test matrix. This search space can often be reduced through knowledge of: the satellite almanac data and current user position and time estimates. This will indicate which satellites/frequencies should be tested first as an attempt to reduce acquisition times, but will not improve an exhaustive search as it does not eliminate any of the search space rather it provides a good initial estimate. This technique has been implemented and tested successfully in software. Results on the collected data set will be presented following the discussion of all applied acquisition methods.

Two published improvements on the serial search technique have been implemented for use with the software radio. The first parallelizes the frequency search space [10]. In this case the raw data is multiplied by the

spreading code with a code phase from the test matrix, then the Fourier transform of the resulting data set is taken. All possible frequency bins are checked for the resulting carrier modulated only with the navigation data which, if found, would indicate the proper code phase had been utilized and the bin in which it was located would provide the necessary frequency information. There is no longer a need to search the various frequency entries of the test matrix, however, the computational requirements of the Fourier transform is substituted for the reduced search space.

The second technique parallelizes the code phase search also through the use of the Fourier transform [11]. The raw data is converted to in-phase and quadrature baseband components using a frequency entry from the test matrix. The Fourier transform of this data is taken and multiplied by the complex conjugate of the Fourier transform of the spreading code. The inverse Fourier transform is then applied to revert back to the time domain. Since multiplication in the frequency domain acts as convolution in the time domain, the resulting data represents the circular convolution at all possible code phases for that particular frequency. Although this requires the computation of the complex Fourier and inverse Fourier transforms, the search space is reduced to only the possible frequency bins. Since the almanac data can often provide optimal frequency starting points, this technique can significantly decrease acquisition times.

The digital correlator/accumulator, popular in the great majority of receivers, make these acquisition techniques impractical since there is no access to the data prior to accumulation. This limitation illustrates the advantage of the software radio approach. Each of the algorithms were coded and tested with the raw data sets displayed in Figure 2. Although each of the algorithms correctly identified the acquisition parameters for all of the satellites in both data sets, the parallelized code phase search technique greatly reduced the exhaustive search times.

The ability to postprocess the raw data provides interesting plots that give deeper insight into the acquisition process. Figure 3 shows the Fourier transform of the GPS-SPS data depicted in Figure 2 post multiplication with the correct C/A code with the proper code phase. This removes the spreading code and the resulting carrier modulated only with navigation data appears at the appropriate frequency. Figure 4 depicts the

same results for data from both GLONASS frequencies. It is important to note that the acquisition signal processing software implemented is applicable to either GNSS. Figure 5 depicts the exhaustive acquisition search results for all entries of the test matrix for a single C/A code from a visible satellite.

The digital correlator/accumulator, popular in the great majority of receivers, make these acquisition techniques impractical since there is no access to the data prior to accumulation. This limitation illustrates the advantage of the software radio approach. Each of the algorithms were coded and tested with the raw data sets displayed in Figure 2. Although each of the algorithms correctly identified the acquisition parameters for all of the satellites in both data sets, the parallelized code phase search technique greatly reduced the exhaustive search times.

Figure 3. Post Acquisition FFT of Raw GPS-SPS Data for All Five Satellite PRN (C/A) Codes

The ability to postprocess the raw data provides interesting plots that give deeper insight into the acquisition process. Figure 3 shows the Fourier transform of the GPS-SPS data depicted in Figure 2 post multiplication with the correct C/A code with the proper code phase. This removes the spreading code and the resulting carrier modulated only with navigation data appears at the appropriate frequency. Figure 4 depicts the same results for data from both GLONASS frequencies. It is important to note that the acquisition signal processing software implemented is applicable to either GNSS. Figure 5 depicts the exhaustive acquisition search results for all entries of the test matrix for a single C/A code from a visible satellite.

SUMMARY

This paper has presented the initial phase in the development of a GNSS software radio. The advantages of such an implementation to the navigation community were discussed along with potential development

Figure 4. Post Acquisition FFT of Raw GLONASS Data for Frequency Channels #21 & #22
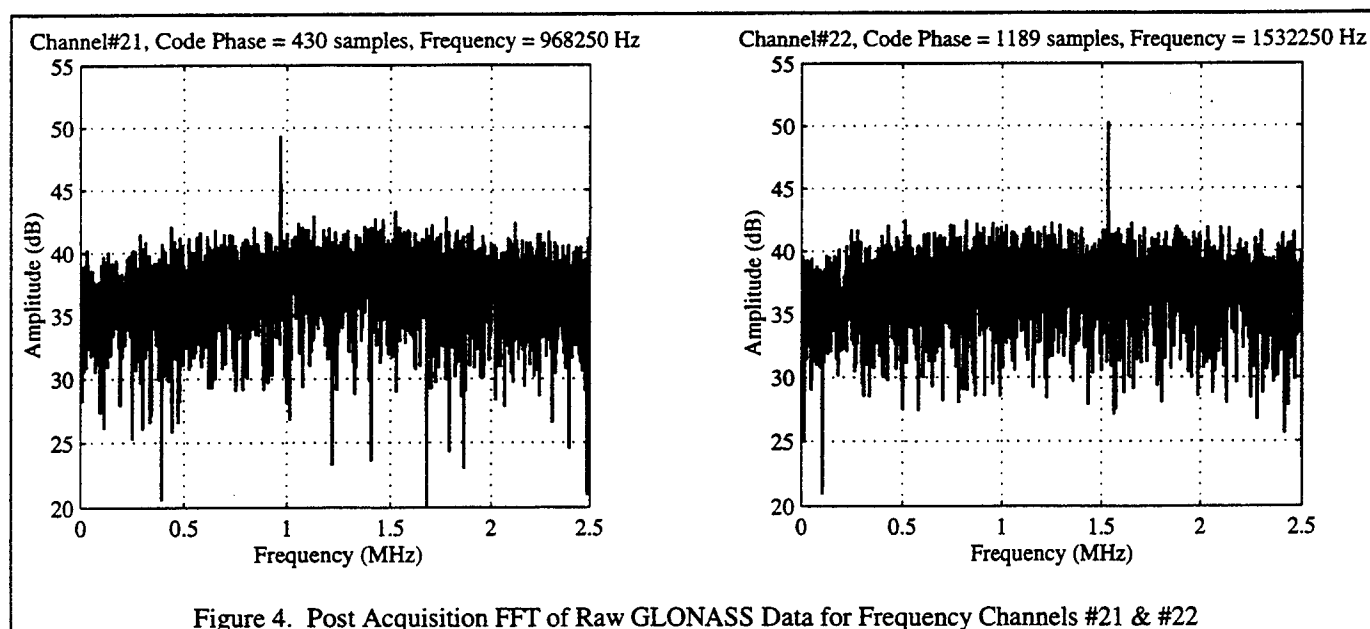
obstacles. To date, a software radio front end has been evaluated and various signal acquisition algorithms, including those which are not applicable in tradition GNSS receiver designs, have been implemented and tested successfully.

Currently, software algorithm implementation of the necessary tracking loops and navigation data processing is underway. Similar to the development of the acquisition algorithms, the tracking algorithms will consist of the traditional implementations as well as any novel innovations developed while at Wright Laboratory [12]. In parallel with the software development is the continued testing of a true GNSS software radio front end. Upon completion of these tasks, efforts will be targeted toward code optimization to finalize the development of the first fully programmable GNSS software radio.

ACKNOWLEDGMENTS

REFERENCES

[1] Vaughan, Rodney G., Scoot, Neil L. and White, D. Rod, "The Theory of Bandpass Sampling", IEEE Transactions on Signal Processing, Vol. 39, No. 9, September 1991, pp. 1973-1984.

[2] Akos, Dennis M., and Tsui, James B. Y., "Design and Implementation of a Direct Digitization GPS Receiver Front End," IEEE Transactions on Microwave Theory and Techniques, Dec. 1996.

Figure 5. Results of Acquisition Algorithm Applied to GPS-SPS Signal over Entire Search Space
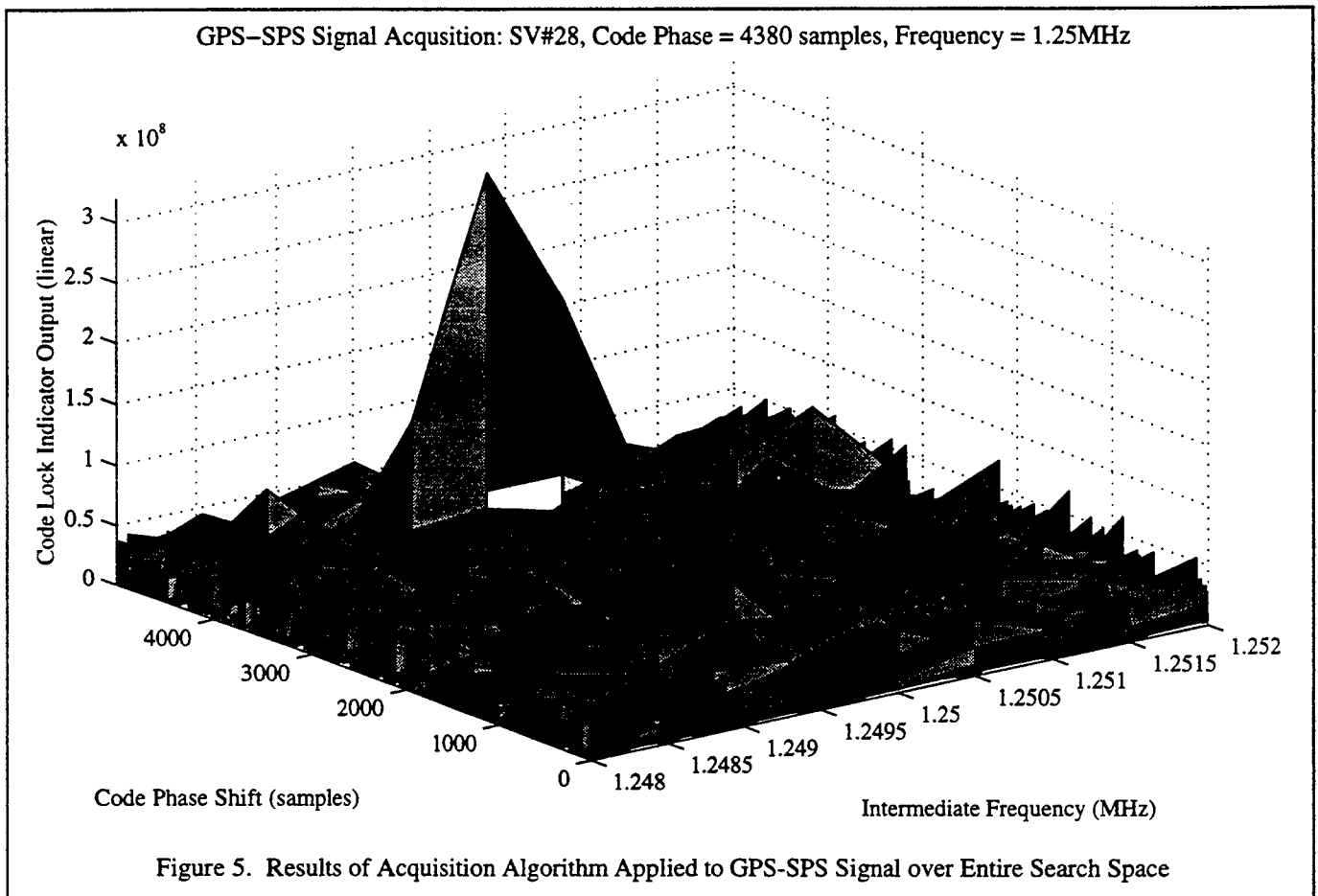
[3] Mitola, Joe, "The Software Radio Architecture," IEEE Communications Magazine, May 1995, Vol. 33, No. 5, pp. 26-38.

[4] Brown, Alison and Wolt, Barry, "Digital L-Band Receiver Architecture with Direct RF Sampling," IEEE 1994 Position Location and Navigation Symposium, Las Vegas, Nevada, April 11-15, 1994, pp. 209-215.

[5] Wepman Jeffery A., "Analog-to-Digital Converters and Their Application in Radio Receivers," IEEE Communications Magazine, May 1995, Vol. 33, No. 5, pp. 39-45.

[6] Baines, Rupert, "The DSP Bottleneck," IEEE Communications Magazine, May 1995, Vol. 33, No. 5, pp. 46-55.

[7] Akos, Dennis M., and Braasch, Michael S.," A Software Radio Approach to Global Navigation Satellite System Receiver Design," 1996 Institute of Navigation Annual Meeting, Cambridge, MA, June 1996

[8] Rappaport, Stephan S., and Grieco, Donald M., "Spread-Spectrum Signal Acquisition: Methods and Technology," IEEE Communications Magazine, June 1984, Vol. 22, No. 6, pp. 6-21.

[9] Holmes, Jack K., and Chen, Chang C., "Acquisition Time Performance of PN Spread-Spectrum Systems," IEEE Transactions on Communications, Aug. 1977, Vol. 25, No. 8, pp. 778-783.

[10] Cheng, Unjeng, Hurd, William J. and Statman, Joseph I., "Spread-Spectrum Code Acquisition in the Presence of Doppler Shift and Data Modulation," IEEE Transactions on Communications, Vol. 38, No. 2, February 1990, pp. 241-250.

[11] van Nee, D. J. R. and Coenen, A. J. R. M., "New Fast GPS Code-Acquisition Technique Using FFT," Electronic Letters, January 17, 1991, Vol. 27, No. 2, pp. 158-160.

[12] Tsui, J. B. Y., Akos, Dennis, and Stockmaster, Michael, "BASS: Block Adjustment of Synchronizing Signal," Wright Laboratories, Patent Pending.

COMPUTER MODELING
OF STRUCTURAL FAILURE

Albert J. Arrieta
Ph.D. Graduate Student
School of Aerospace
and Mechanical Engineering

University of Oklahoma
865 Asp Av
Norman, OK 73019

Final Report for:
Graduate Student Research Program
Wright Laboratory

August 1996

COMPUTER MODELING
OF STRUCTURAL FAILURE

Albert J. Arrieta
Ph.D. Graduate Student
School of Aerospace
and Mechanical Engineering
University of Oklahoma

Abstract

Several computer codes were evaluated that predicted failure from either buckling and/or damage such as cracks and delaminations. The Knowledge Systems Inc.(KSI) BuckDel computer program performs buckling analysis of laminated composite shell structures with delaminations in order to determine the remaining strength. Also, the KSI phase II SBIR proposal was evaluated. This proposal would incorporate damage tolerance analysis software into the Automated Structural Optimization Systems (ASTROS) software. The ASTROS V12 computer program combines mathematical optimization routines with classical structural analysis software and was recently modified to include panel and column buckling capabilities. The ASTROS panel and column buckling analysis performed well when compared to handbook solutions but could only be used separately. The panel and column buckling capabilities in ASTROS could not be used at the same time. Buckling constraints were used in ASTROS to optimize the design of a simple aircraft wing for minimum weight. Later, fatigue crack damage tolerance constraints were applied to the simple aircraft wing and the ASTROS optimization resulted in a weight increase of 49%. Finally a p-element finite element code called Mechanica was evaluated on a two cell beam and on a simple aircraft wing. Static and modal analyses were compared between ASTROS and Mechanica with mixed results.

COMPUTER MODELING
OF STRUCTURAL FAILURE

Albert J. Arrieta

Introduction

The ability to predict failure in structures has always been a goal of engineers and designers. This is especially true when the structure is a ground or flight vehicle which must be light to minimize costs and maximize performance yet strong enough to insure structural integrity and safety throughout its operating life. Unfortunately minimizing weight and maximizing structural strength are opposing objectives. A compromise between these two objectives is often made. The traditional finite element analysis (FEA) computer program provided the engineer and designer with the maximum stresses in the structure which could be compared to the material's allowables. If the calculated stresses were too high, a redesign of the structure would follow. Typically the traditional FEA was performed without considering buckling, fatigue cracks or composite delamination. Buckling failure would be of concern with long and slender components in the structure experiencing compressive loading. Fatigue and composite delamination failure modes can be expected to develop in the structure during its operational life. In light vehicle structures these failure modes are detrimental to structural integrity, costly to repair, compromise safety and will eventually lead to catastrophic failure. The designer or engineer requires a methodology to determine what effect buckling, fatigue, and composite delaminations have on the design of light vehicle structures so as to insure durability and safety at minimum weight. This report reviews three computer programs which could be used in this determination.

2-3

Discussion And Results

1. Buckling & Delamination Analysis of Stiffened Laminated Plates (BuckDel)

The BuckDel computer program was developed by Knowledge Systems Inc. to model the strength reduction due to delaminations and to model the mechanism of delamination growth. This software performs geometrically nonlinear analysis of stiffened laminated composite shells with and without delaminations. It is capable of linear static and linear buckling (eigenvalue) analysis. A nonlinear post-buckling analysis is performed though limit and bifurcation points. BuckDel has two finites elements, a triangular laminated shell element and a fiber-reinforced beam element. All materials modeled must be linearly elastic. The multi-domain modeling method is used for analysis of laminated plates/shells containing delaminations. The model consist of the delaminate, base, and undelaminate. The delaminate and base model the top and bottom of the delamination and designate the boundary (delamination front) between the undamaged parent material(undelaminate) and the delaminated composite. The BuckDel computer program can also provide delamination growth prediction by using pointwise energy release rate at the delamination front. The previous information was taken from the BuckDel v0.9 Users and Theory manual. The software code was not available for evaluation.

2. Evaluation of Buckling Constraints in ASTROS

Version 12.0 of ASTROS advertises a stiffened panel buckling capability. Constraint cards were developed to limit design variables during an optimization that were affected by Euler column buckling or panel buckling loading. The ASTROS documentation states that this version "Considers plate bending and stiffener behavior independently but simultaneously". This statement implies that both column and panel buckling constraints could be applied in the same ASTROS run, but ASTROS only returned errors. The

2-4

following results are from the ASTROS output when the column and panel buckling constraints were applied separately and finally when they were applied simultaneously. The unstiffened panel buckling constraint is driven by plate bending elements with the DCONBK card. The panel is simply supported on its edges regardless of any boundary conditions in the finite element model. The panel buckling constraint, DCONBK, can only be applied to QUAD4 and TRIA3 elements. The panel running loads N_x, N_y, S, and the laminate bending stiffness matrix D are provided by the finite element model. The actual panel buckling analysis are performed on a pseudo-panel which takes its dimensions from the appropriate finite elements or from the user supplied DCONBK card. The user may also supply a safety factor, $\lambda_{all} = N_c/N$ in the DCONBK card. N_c is the critical buckling load vector and N is the finite element model running load vector that is currently applied to the panel. Equation 1 defines the relationship.

$$N_c = (N_{CX}^2 + N_{CY}^2 + S_C^2)^{1/2} = \lambda_{all}(N_X^2 + N_Y^2 + S^2)^{1/2} = \lambda_{all}N \qquad (1)$$

A λ_{all} of 2.0 means that ASTROS will optimize the panel (QUAD4 element) so that the critical buckling load is twice the finite element model load. A buckling test case of an unstiffened panel was performed by ASTROS. A single QUAD4 element was uniaxially loaded with a running compression load of N_x = 8.33 lb_f/in and a safety factor of λ_{all} = 1.0. The QUAD4 element was a model of a rectangular flat plate with length(L), width(w), and thickness(t) of 30.0 inches, 6.0 inches and 0.2 inches, respectively (figure 1).
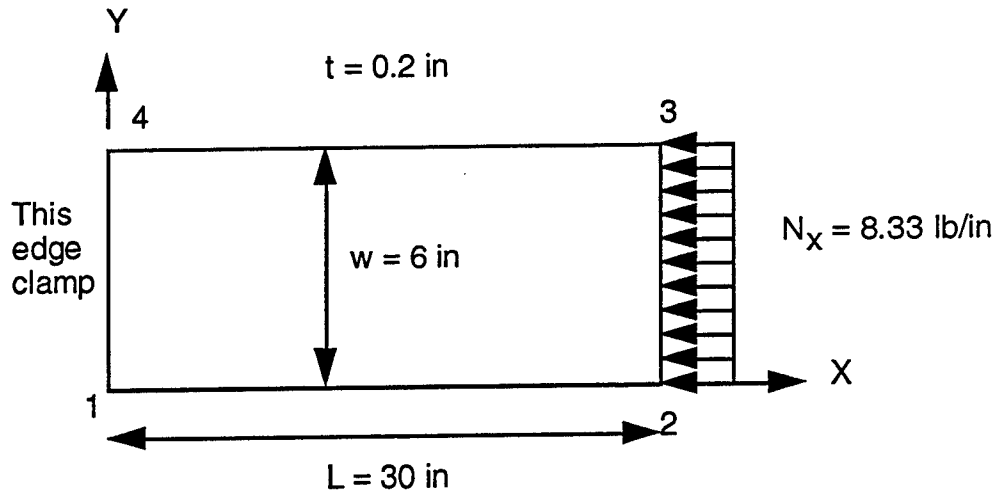
Figure 1: Panel Buckling Test Case Problem

After optimization the thickness was reduced to 0.0202 inch. With this new thickness the handbook buckling calculation (eqn 2) of a panel in uniaxial loading was 8.35 lb_f/in. With an input load of 8.33 lb_f/in ASTROS calculated the correct panel buckling load to within 0.2%. The column buckling

$$N_{cr} = 3.62tE(t/w)^2 \qquad\qquad E = 10x10^6 psi \qquad\qquad (2)$$

analysis are driven by the 1-D bending elements and the constraint card DCONBKE. Column buckling constraints can only be applied to BAR and ROD elements. The column buckling equations are based on Euler buckling equations. Four different end boundary conditions are available for the user to pick from. The end boundary conditions are pinned-pinned, fixed-free, fixed-pinned, and fixed-fixed, and are specified in the DCONBKE card. The user may specify a safety factor, λ_{all} = P_{cr}/p where P_{cr} is the critical Euler buckling load and P is the finite element model applied load. For the test case a single rod element with pin-pin boundary conditions and a safety factor of 2.0 was analyzed with ASTROS. The rod element was 15 inches long and had

an initial cross-sectional area of 1.0 in². After optimizing with a 50 lb$_f$ uniaxial compressive load, the final cross-sectional area was 0.0535 in². Using this final area (A) in the theoretical Euler buckling equation (eqn 3) predicted a critical buckling load of 100 lb$_f$ which was twice the applied load.

$$P_{cr} = \frac{\pi^2 EI}{L^2} = \frac{\pi EA^2}{4L^2} \qquad (3)$$

The two test cases performed for this report indicate that ASTROS's buckling capabilities appear to be accurate for simple problems. During these ASTROS analyses, two problems were encountered when the buckling constraints were employed. With the Convex version of ASTROS (v12) using the DCONBK panel buckling constraint, the following message was returned: "USER FATAL MESSAGE, NUMBER 12.4, NO CONSTRAINTS HAVE BEEN APPLIED IN A DESIGN OPTIMIZATION RUN". Then ASTROS terminated with errors during the first iteration. The optimization problem was the same single QUAD4 element under compression loading model used in the previous panel buckling test case. This same model ran properly on the IBM ASTROS_UAI version 12. Another problem exist with the fully stressed design (FSD) optimization strategy in ASTROS_UAI v12 running on the IBM computer. ASTROS returned a "DATABASE FATAL ERROR DBOPEN01 ENCOUNTERED FOR ENTITY NAME CONST IN ROUTINE DBOPEN". Then ASTROS terminated with errors during the first iteration. No problem occurred when the MP optimization strategy was used. The optimization problem was a single ROD element under compression with a DCONBKE column buckling constraint. This same problem ran properly on the CONVEX ASTROS v12. The next logical step was to apply both column and panel buckling constraints in the same ASTROS run on the IBM ASTROS_UAI version. This test case used the same rectangular panel in the previous panel buckling test case. This time two horizontal stiffeners

were added to the top and bottom edge. The panel was modeled with a QUAD4 and the stiffeners with RODs and five different ASTROS runs were made. All three elements were designed and the same uniaxial running load of N_x=8.33 lb/in was applied as in the previous panel buckling test case. For the first run only column buckling constraints were applied to the two ROD elements. The problems converged with a new weight of 0.67 lb. In the second run only panel buckling constraints were applied on the single QUAD4 element and this problem converged with a new weight of 0.89 lb. The third run used column buckling constraints on the ROD elements and a panel buckling constraint on the QUAD4 element. With the same DCON SID = 1 in the solution control deck to call both DCONBKE and DCONBK cards, ASTROS did not run and returned the statement "ERROR MMGETB 001-BLOCK UNT1 ALREADY ALLOCATED". The fourth run again used both types of buckling constraints as before but used two different DCON SIDs under the single boundary condition set in solution control. This solution control follows.

```
                BOUNDARY SPC = 1
                        STATICS(MECH = 1,DCON = 1)
                        STATICS(MECH = 1,DCON = 2)
                END
```

ASTROS failed to run and returned "MMGETB 001-BLOCK UNT1 ALREADY ALLOCATED". The fifth and final run used both types of buckling constraints as before but used two different DCON SIDs under two different boundary condition sets in solution control as presented below.

```
                BOUNDARY SPC = 1
                        STATICS(MECH = 1,DCON = 1)
                BOUNDARY SPC = 1
                        STATICS(MECH = 1,DCON = 2)
                END
```

Again ASTROS failed to run and returned the same error message. The conclusion made is that panel buckling and column buckling constraints cannot be used together in this version of ASTROS.

3. AME 6763 Simple Aircraft Wing With Buckling Constraints

The AME 6763 simple aircraft wing is a finite element model used in the University of Oklahoma Aerospace and Mechanical Engineering(AME) School Advanced Finite Element course AME 6763. This model is a three spar, five bay wing 10 ft by 20 ft, made of 113 elements and a 12,000 lb_f wing tip load. The spar and rib webs were modeled with SHEAR elements. The spar and rib caps were modeled with ROD elements. The original weight was 538.1 lb. After an ASTROS optimization with only Von Mises stress constraints, the new weight was 318.3 lb. Panel buckling constraints were added to the upper wing skin elements and the wing was optimized with a buckling safety factor $\lambda_{all}=1.3$. With panel buckling constraints and Von Mises stress constraints applied, the new wing weight was increased to 1,536.3 lb by ASTROS. For the second run column buckling constraints were applied to the upper spar caps along with Von Mises stress constraints for the whole wing. With a column buckling safety factor of $\lambda_{all}=1.3$, ASTROS returned a new wing weight of 654.3 lb. Apparently panel buckling constraints are more severe than column buckling constraints for achieving minimum mass structures.

4. Damage Tolerance in An Optimization Environment

The stress and strain constraints available in ASTROS are used to meet only static strength requirements. These constraints do not consider long term durability from other failure modes such as fatigue and corrosion. In industry another requirement imposed on new aircraft design and repairs is a damage tolerant requirement. This requirement insures the structural integrity or residual strength of a structure due to the presence of a hypothetical sharp crack growing in the worst orientation in the most critical members. Typically damage tolerance requirements specify that the cracked structure retain sufficient residual strength for a certain length of time or

operational usage such as flight hours. In this research a simple damage tolerant requirement of 100,000 cycles were imposed on the AME 6763 aircraft wing through global/local analyses. The AME 6763 wing was considered the global finite element model. From previous ASTROS runs the maximum tensile stresses in the lower wing skin were found in QDMEM1 element number 8. A separate local finite element model of this location represented by element 8 was created with a 2x2 mesh of QUAD4 elements. The grid displacements from the global FEA model of the wing were used as boundary conditions for the local finite element model. Linear interpolation of the local finite element model boundary conditions was necessary for the intermediate nodes. For simplicity, a constant amplitude fatigue loading and a wide plate through crack stress intensity factor (eqn 4) were assumed.

$$K_I = \sigma\sqrt{\pi a} \tag{4}$$

$$N_f = \frac{1}{C\sigma_{max}^n \pi^{\frac{n}{2}}[1-R]^{mn}} \int_{a_i}^{a_f} \frac{da}{a^{\frac{n}{2}}} \quad , C = 1x10^{-9}, R = 0.1, m = 3, n = 4 \tag{5}$$

The Walker crack growth equation and the stress intensity factor in equation 4 were used to calculate the peak cyclic stress element 8 could sustain for 100,000 cycles of constant amplitude loading. From equation 5, the peak cyclic stress was computed as 10.89 ksi. This peak cyclic stress was divided by the stress concentration factor K_t derived from the local finite element model. This new value of peak cyclic stress was set as a maximum principal stress constraint with the ASTROS function package. ASTROS optimization runs were performed on the AME 6763 wing global finite element model with a maximum principal stress constraint. The first ASTROS run ignored the stress concentration factor computed by the local finite element model. With a maximum principal stress constraint of 10,890 psi, the AME 6763 wing increase

2-10

in weight by 29% from 538.8 lb to 696.2 lb. The second ASTROS run included the stress concentration factor computed by the local finite element model. With the maximum principal stress constraint reduced by K_t to 9,595 psi, the AME 6763 wing increased in weight by 49% from 538.8 lb to 801.3 lb. For the third run the local finite element model K_t was ignored and a maximum principal stress constraint of 10,890 psi along with upper skin buckling constraints ($\lambda_{all}=1.3$) were used by ASTROS. The AME 6763 wing increased in weight by 224% from 538.8 lb to 1,748 lb. The fourth and final run used the local finite element model K_t to reduce the maximum principal stress constraint to 9,590 psi. Along with upper skin buckling constraint ASTROS increased the AME 6763 wing weight by 223% from 538.8 lb to 1,796 lb. The small difference in final weight between runs 3 and 4 indicated that the buckling constraint had a dominating effect on the optimization.

5. Mechanica Finite Element Software Review

Mechanica is a finite element software program developed by the Rasna Corporation. Rasna was later purchased by Parametric Technology. Mechanica uses a 2-D or 3-D p-element optimization code. Its element library contains the usual finite elements such as beams, thin shells and solids. Mechanica does not have SHEAR or ROD elements. Mechanica uses the p-element method which increases the order of the elements during remeshing instead of increasing the number of elements. Each remeshing is called a P-loop pass or iteration. Mechanica continues to increase the order of elements in the model until a user specified convergence criterion is met. This criterion is typically met when the values of the strain energy and displacement of the previous P loop pass is within 10% of the current values. The task of this evaluation was to perform static and modal analysis on built up structures. The first finite element model was a two cell cantilever box beam shown in

figure 2. The box beam had a length of 40 inches, a width of 20 inches and a depth of 10 inches. This model had three spars with webs modeled by thin shell elements and caps modeled by beam elements with circular cross-sections. The outer surface of the box beam consisted of thin shell elements. All thin shell elements were 0.1 inch thick, and the cross-sectional areas of the beam elements were 1.0 in^2. One end of the box beam was fixed in all six degrees of freedom which is indicated by the triangles. The other end was loaded by 300 pounds distributed equally among 6 gridpoints. Only 104 elements were needed for the finite element solution to converge. The initial static stress analysis did not include spar caps but was added on subsequent analyses. For static stress analysis, the strain energy and displacement converged in 4 P-loop passes. The Von Mises, maximum principal, and ZZ-normal stresses converged as well. The maximum bending stress, without spar caps in the model, appeared to be between 400 psi and 538 psi. This range of maximum bending stresses compared favorably with the theoretical beam bending stress computed in equation 6.

$$\sigma_B = \frac{My}{I_x} = \frac{(12,000 in \cdot lb) \times (5.0 in)}{121.54 in^4} = 498.6 psi \tag{6}$$

The second static stress analysis with Mechanica included spar caps which reduced the maximum bending stresses in the bottom skin to 161-215 psi. The Von Mises, maximum principal, and ZZ-normal stresses converged in 4 P-loop passes. The subsequent modal analysis converged with a mode 1 natural frequency of 142 Hz.

6. Comparison Between ASTROS And Mechanica

Static stress and modal analyses were performed on the AME 6763 aircraft wing with ASTROS and Mechanica. The maximum principal stress, natural

frequency and mode shapes computed by ASTROS and Mechanica were compared. The ASTROS finite element model of the AME 6763 wing is the same model discussed in section 4. The Mechanica finite element model of the AME 6763 aircraft wing was slightly different than the ASTROS model. Mechanica does not have shear, membrane or rod elements. Therefore Mechanica modeled the webs and skins with thin shells and modeled the spar and rib caps with beam elements. The Mechanica static stress analysis converged and a maximum principal stress contour plot of the lower wing skin is shown in figure 3. The maximum principal stresses are significantly higher than ASTROS results. More disturbing was the discontinuous stress distribution in the skins across the ribs. The results from the static analysis are listed in Table 1. The Mechanica modal analysis of the AME 6763 wing converged in two P-loop passes. The first three mode shapes are displayed in figures 4, 5, and 6, respectively. The Mechanica mode one natural frequency approximated the ASTROS mode one frequency. Accordingly, this mode shape was the same as the ASTROS results. The Mechanica mode 2 and 3 frequencies were higher than the ASTROS results. This should be expected because the Mechanica finite element model was stiffer than the ASTROS model. The Mechanica model was stiffer because it used thin shell elements instead of shear and membrane elements, and beam elements instead of rod elements. The results from the modal analysis are in Table 2.

Table 1: Static Analysis Comparison

	ASTROS	Mechanica
Maximum Principal Stress (psi)	20,385	27,486
Maximum Displacement (in)	13.7	9.84

Table 2: Modal Analysis Comparison

	Natural Frequency (Hz)	Mode Shape
ASTROS		
MODE 1	12.1	Bending
MODE 2	39.7	Torsion
MODE 3	50.0	Torsion & Bending
Mechanica		
MODE 1	12.7	Bending
MODE 2	45.8	Torsion & Bending
MODE 3	54.7	Torsion & Bending
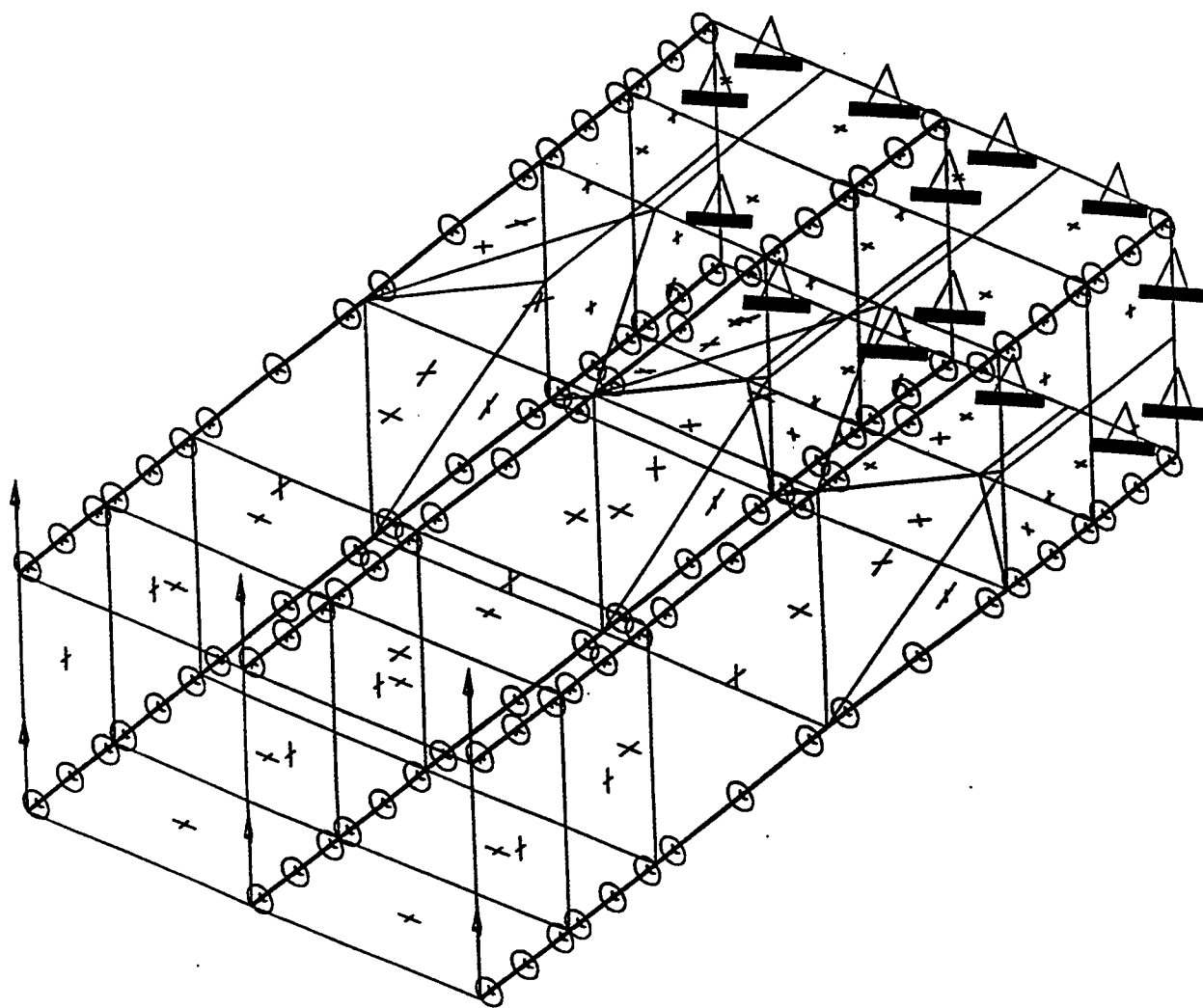
Figure 2: Mechanica Finite Element Model of Box Beam
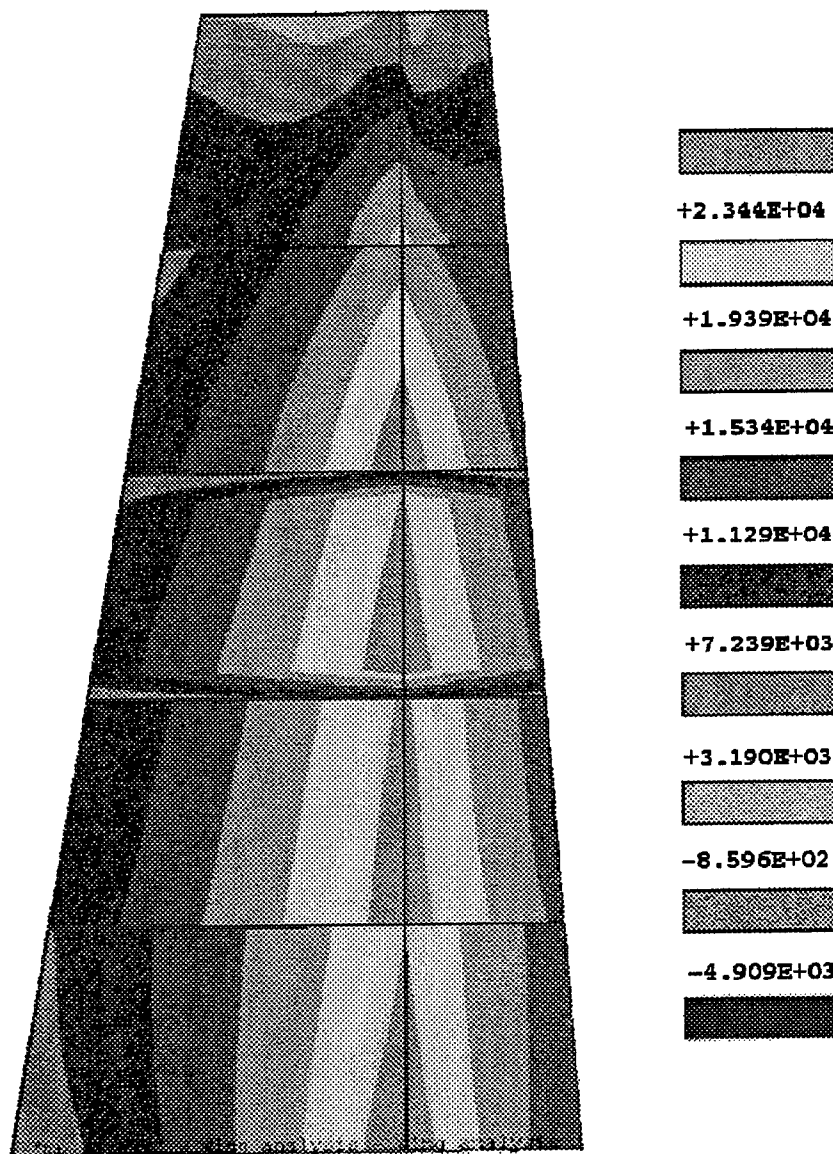
Figure 3: AME 6763 Wing Lower Skin Stress Contour

The legend values, top to bottom:

+2.344E+04

+1.939E+04

+1.534E+04

+1.129E+04

+7.239E+03

+3.190E+03

-8.596E+02

-4.909E+03
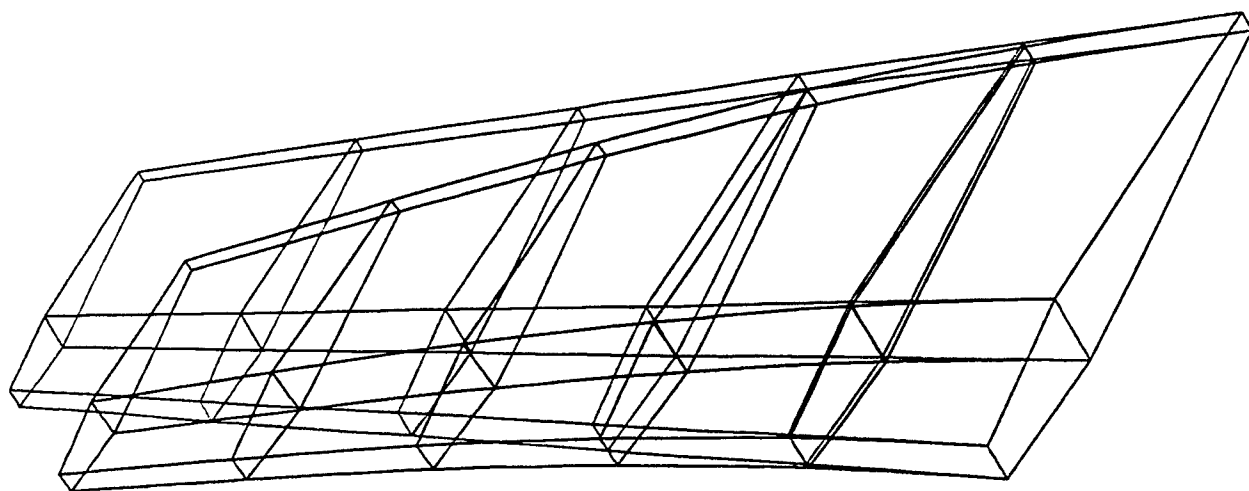
Figure 4: AME 6763 Wing, Mode Shape 1
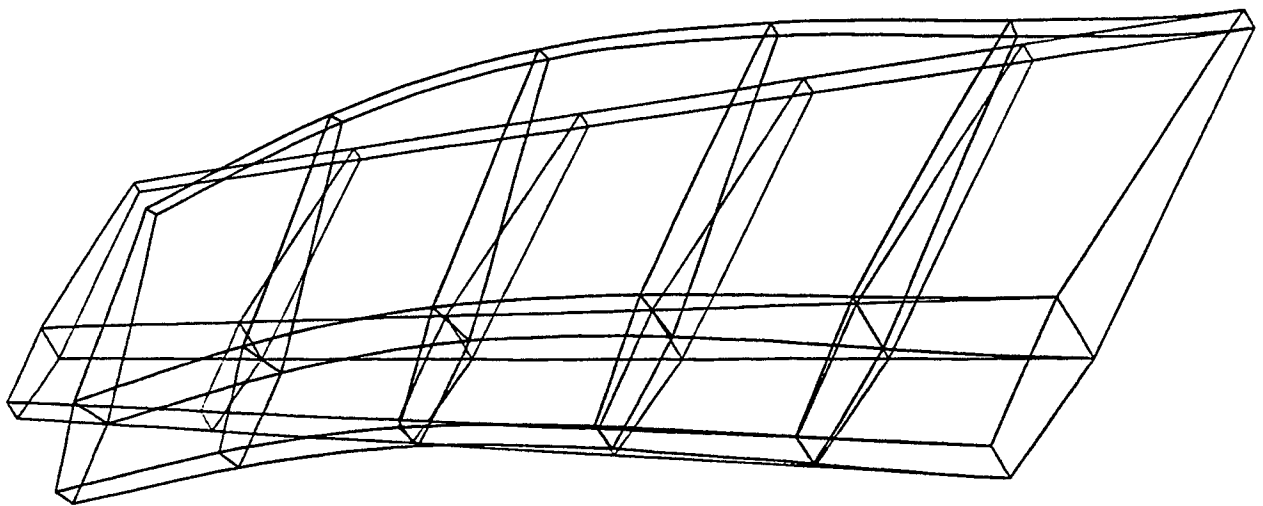
Figure 5: AME 6763 Wing, Mode Shape 2
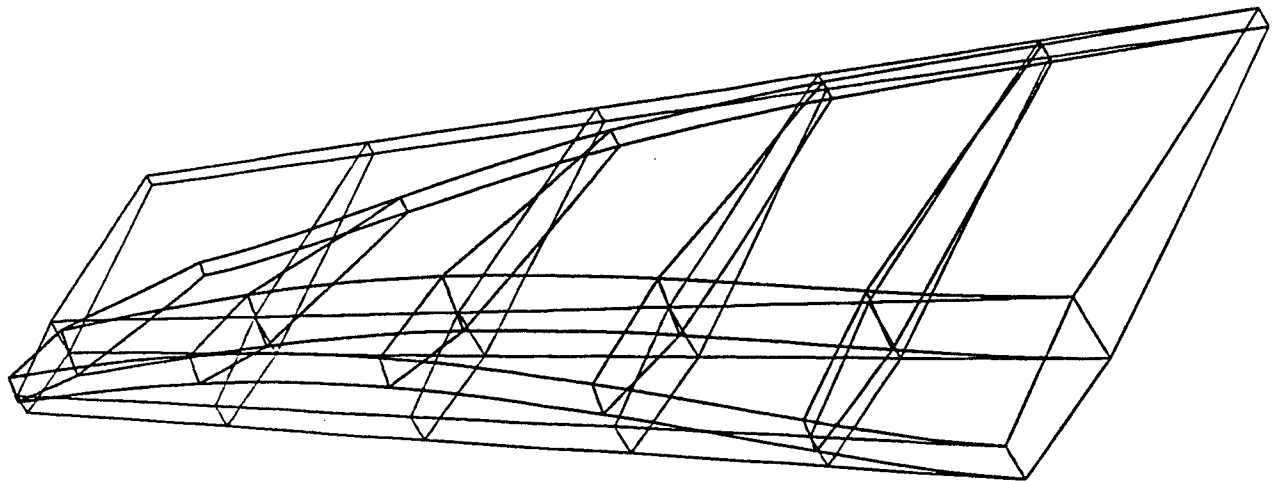
Figure 6: AME 6763 Wing, Mode Shape 3

A FAST FOURIER TRANSFORM ANALYSIS
OF PILOT INDUCED OSCILLATIONS

Sten E. Berge
Graduate Student

School of Aeronautics and Astronautics
Purdue University
West Lafayette, IN 47907-1282

A FAST FOURIER TRANSFORM ANALYSIS
OF PILOT INDUCED OSCILLATIONS

Dominick Andrisani, II
Associate Professor

Sten E. Berge
Graduate Student

School of Aeronautics and Astronautics
Purdue University

Abstract

During the summer of 1996 a program was initiated by the authors to study the characteristics of pilot induced oscillations (PIOs). The long term goals are

- to develop a methodology to take time history data from flight or ground based simulation and determine if a vehicle is PIO prone,

- to determine if a PIO actually occurred in a time history record,

- to develop a non-real-time analysis tool to determine from flight simulator time history data if PIOs occurred so that simulation engineers can help insure consistency between pilot comments, Cooper-Harper ratings, PIO ratings and time history data, and

- to develop a real-time capability of detecting a PIO fast enough to take action to prevent the full development of the PIO.

Results obtained over the period 7/1/96-8/23/96 deal primarily with the first goal and are described in this report. A computer analysis tool is developed in which fast fourier transforms are used to determine

- resonant frequency and output phase angles at the resonant frequency,

- predictions of PIO susceptibility using the Smith-Geddes PIO criteria.

The following preliminary results are highly encouraging.

- 13 of 14 configurations which experienced PIOs were correctly identified.

- In 12 additional configurations 6 were correctly identified as being not PIO prone. Results for the remaining 6 configurations were complicated by multiple resonances that confuse the resonance detector. However, depending on the choice of resonant frequency the correct classification is made for each of the 6 configurations.

- The occurrence of multiple resonances is closely associated with configurations that are not PIO prone.

- continued development of the resonance detector is planned.

A FAST FOURIER TRANSFORM ANALYSIS
OF PILOT INDUCED OSCILLATIONS

Dominick Andrisani, II and Sten E. Berge

Background

Pilot induced oscillations (PIOs) have received much recent attention [2-6]. Several approaches to development of a unified PIO theory are currently being pursued. A catalog of PIO events has been compiled and several existing criteria have been shown to successfully predict PIO susceptibility. Accurate analysis of the ever-growing PIO database will gather insight into the characteristics of PIOs. With this information, a PIO detector can be designed for time history data as generated in either flight or simulation.

Discretization of Time History Data

Most PIO time history data in the catalog of PIO events is in paper copy form; the original flight data having been lost or unavailable for various reasons. As a first step in this analysis, various time history records from the PIO database were digitized. Eleven classic PIOs and eighteen Bjorkman (NT-33) time histories [7] were digitized. The eleven classic PIOs were taken from the PIO catalog compiled by HAI in cooperation with STI[2]. There is one roll ratchet, several rate limited actuator cases, mode switches, and both lateral and longitudinal examples. The Bjorkman time histories are all longitudinal, but provide a consistent database encompassing nearly the entire range of the PIO tendency classification scale. The digitization was hampered by the quality of the reproductions and the scales of the plots. Sections of data were often missing because of repeated previous reproduction. Digitization noise was introduced into the data axis scaling was not well chosen. For example, the B-2 charts cover too much time, and the non-PIO Bjorkman configurations are plotted with the same scale as the large amplitude PIO histories. Conversely, there were several data sets with time scaling that covered too small a time period and which cut off the peaks of the data. The M2F2 charts are examples of this other extreme.

The following multi-step digitization process was employed.

- PIO time histories were scanned and cleaned up using OFOTO commercial software.
- PIO time histories were digitized the data using dataThief shareware.
- The resulting ASCII data file was converted to MATLAB script format.

- A MATLAB script was written to produce a second MATLAB script file with a specified uniform time increment.
- A MATLAB script was written to plot the time history data in a similar format to the original.

In all 27 longitudinal records (flight records from the T-38, Shuttle, B-2, F-8, YF-22, and X-15, 18 Bjorkman NT-33 flight records, and 3 HAVE PIO ground based simulation records) and 5 lateral records (flight records from the F/A-18(2), F-14, YF-16, and M2-F2) were digitized. These records are available at WL/FIGC-2 (Mr. David Leggett is the point of contact, 513-255-8498).

Technical Approach

A spectral analysis is performed on selected time histories to determine the quantities needed for PIO characterization. The resonance of each measurement is determined from power spectral density for that variable. The actual resonant frequency will differ from the discrete frequencies used in the power spectral density plot, unless the resonance happens to correspond to one of the harmonics of the spectral analysis. Therefore, the resonant frequency is approximated by finding the centroid of the peak and its two nearest neighbors. The cross spectrum between input and output is used to calculate a frequency response function relating input and output. The phase angle at the resonant frequency is taken from the phase angle of the frequency response function. The validity of the results is related to the coherence of the cross spectral data points used.

The Smith-Geddes criteria are then evaluated. The pilot-vehicle resonant frequency is chosen to be the resonant frequency of the variable with the most reliable time history. This is most often pitch rate. The Smith-Geddes criteria require the phase angles of both $\Theta(j\omega)/F_s(j\omega)$ and $N_z(j\omega)/Fs(j\omega)$. The phase of $\Theta(j\omega)/F_s(j\omega)$ can also be calculated by subtracting 90 degrees from the phase of $q(j\omega)/F_s(j\omega)$. The Smith-Geddes analysis is done with both the Θ and q phase angles. Susceptibility to PIO as defined in MIL-STD-1797A [8], approximate Cooper-Harper rating, and approximate PIO rating based on Ralph Smith's data method [10] are outputs of this analysis.

This procedure is not be completely automatic at this time for two reasons. First, the FFT window size must be chosen subject to several constraints: the primary (lowest) harmonic (set by the length of time (T_{window}) used in the FFT, $\Delta\omega=2\pi/T_{window}$) should be approximately 1 rad/sec, and the highest harmonic (set by the time increment (ΔT), $\omega_{max}=\pi/\Delta T$) must be at least 20 rad/sec to include the resonant frequencies of all possible types of PIOs including pitch bobbles and roll ratchets, the window not contain discrete events such as control system mode switches, and at least two similar windows are required

to compute coherence. Second, the best pilot-vehicle resonant frequency must be manually chosen when multiple resonances are present. Unless the same dominant resonant frequency appears in all measurements, the correct choice is unclear. Further research on this problem is continuing.

The spectral analysis is performed by MATLAB's spectrum function in the Signal Processing Toolbox. The procedure used is Welch's averaged periodogram method [1,9]. Each frame consists of three subframes (two consecutive and one overlapping one half of the other two). Each subframe is preprocessed by removing any bias and is windowed with a Hanning window. Use of more than one subframe helps insure valid coherence. The software steps through the time history frame by frame.

PIO susceptibility is determined by the Smith-Geddes criteria using the following steps.

- Pilot-in-the-Loop Properties

    Resonant frequency, wr,

    Amplitudes of inputs and outputs at wr from FFT data

$$\Theta(j\omega_i), \quad q(j\omega_i), \quad n_z(j\omega_i).$$

- Pilot-Vehicle System Properties: Phase lags computed using Cross Spectral densities

$$\angle\left.\frac{\Theta(j\omega)}{F_s(j\omega)}\right|_{w=wr} \qquad \angle\left.\frac{q(j\omega)}{F_s(j\omega)}\right|_{w=wr} \qquad \angle\left.\frac{n_z(j\omega)}{F_s(j\omega)}\right|_{w=wr}.$$

- Smith-Geddes PIO Metrics

    Use wr as the criterion frequency,

    Determine susceptibility to

        attitude type PIO (Type III) based on

$$\angle\left.\frac{\Theta(j\omega)}{F_s(j\omega)}\right|_{w=wr} < -180 \qquad \angle\left.\frac{q(j\omega)}{F_s(j\omega)}\right|_{w=wr} -90 < -180,$$

        acceleration type PIO (Type I) based on

$$\Phi(\omega_r) = \angle\left.\frac{n_z(j\omega)}{F_s(j\omega)}\right|_{w=wr} - 14.3\omega_r < -180 \text{ deg. when}$$

$$\angle\left.\frac{\Theta(j\omega)}{F_s(j\omega)}\right|_{w=wr} < -160$$

A few comments are in order. We evaluate phase angles at the resonant frequency of the pilot-in-the-loop system, w_r. The Smith-Geddes criterion evaluates the phase angles at the criterion frequency, $w_c = 0.24S + 6$ where S is the average slope of the $|q(j\omega)/Fs(j\omega)|$ over the interval from 1-6 radians per second. wc is an approximation to the crossover frequency of the pilot-vehicle system during pitch attitude tracking. If there is a dominant resonance in a time history record our approach is reasonable since the pilot-vehicle system will resonate near the crossover frequency of the pilot-vehicle system.

Our approach is more than an open loop method because our resonant frequency is a property of the pilot-vehicle system as expressed in the time histories from the piloted system.

Problems with the Resonant Frequency Detector

Several problems were encountered with our resonant frequency detector.

• Power spectra of important traces sometimes have multiple peaks. Figure 1 shows an example of a time spectrum with two resonant peaks. Figure 2 shows the more common case where there is a single dominant resonant peak.

• Power spectra of different traces sometimes have different peaks as shown in the sketch below.
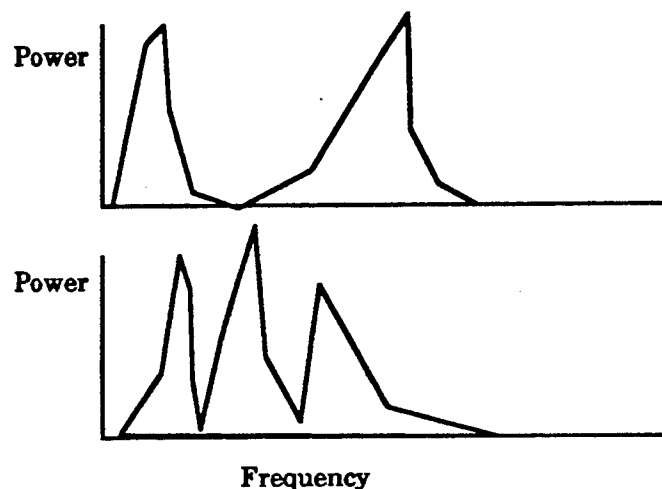
• Low frequency power sometimes confuses our resonant frequency detector. Figure 1 also illustrates this problem.

There are several possible solutions to those problems. Multiple peaks on one power spectrum and peaks that don't line up seem to correlate with the absence of PIO. Low frequency power might be removed by a different choice of windowing in computing the FFT

or by filtering the time history before the FFT is taken. We are currently using a Hanning window.

The end of the summer prevented our resolution of the problems. Development of the resonant frequency detector will continue until the end of 1996 by Professor Andrisani.

Results

The results from our analysis of longitudinal PIO data are summarized in Tables 1 and 2.

PIO amplitudes versus resonant frequency is shown in Figure 3. The longitudinal PIOs in this study occur at a wide range of amplitudes. This suggests amplitude of the resonance is not a good metric to determine if a PIO is occurring.

Angle Φ (of the Smith-Geddes criterion [10]) versus actual average PIO rating for the Bjorkman data is shown in Figure 4. The line in the figure comes from [10] and seems to underestimate the average PIO ratings. Average Cooper-Harper ratings versus angle of $\Theta(j\omega)/Fs(j\omega)$ is shown in Figure 5.

Figure 6 shows that the occurrence of a PIO seems to correlate well with a single dominant resonant frequency in the time history. This characteristic seems more helpful in determining that a PIO has occurred than output amplitude at the resonant frequency. More work is required here to determine how to effectively use this fact.

Table 1 indicates that acceleration type PIO were important in detecting PIO susceptibility in only 1 of 24 cases (B2). The need for acceleration type PIO is currently under review by the authors. These results indicate that the Smith-Geddes boundaries are reasonable.

When phase angles vary significantly during a time history nonlinear behavior is suspected. One cause of this is rate limiting, where phase lag is a function of amplitude and frequency. The Shuttle record, Figure 7, shows strong time varying phase angle of $q(j\omega)/Fs(j\omega)$. The B2 record, Figure 8, shows almost no time varying phase angle of $q(j\omega)/Fs(j\omega)$. The ability of this method to track time varying phase angles is significant because this can trigger a PIO.

Towards a Time Domain PIO Detector

The following comments summarize the status of our efforts at developing a FFT Based PIO detector.

- The existing algorithm is not set up to run in real time. A Wright Laboratory in-house effort by Mr. Phil McKeehan is anticipated in this area.

- Some operator intervention is helpful in selecting the time periods to analyze so that PIO events are not split into different time frames.
- Improvement is necessary on the resonant frequency detector.
- We still need to determine in what way amplitude at resonant frequency is helpful for PIO detection.
- We need to decide what the occurrence of multiple resonances indicates about PIO.

Conclusions

Based on recently digitized time histories of 27 longitudinal PIOs the following conclusions are drawn.

- Our FFT Based PIO Analysis Software using the Smith-Geddes criteria does an excellent job of determining PIO susceptibility.
- Preliminary results indicate that the Smith-Geddes boundaries are reasonable.
- Acceleration type PIO were important in detecting PIO proneness in only 1 of 24 cases (B2). The need for this characterization is under continuing review.
- Continued development of the resonant frequency detector is required for cases where multiple resonances occur.
- The ability of this method to track time varying phase angles is significant because this can serve as a trigger for PIOs.

Further Work

During Professor Andrisani's sabbatical at Wright Laboratories (8/19/96-12/20/96) the additional work on the following items is planned.

- resonant frequency detector,
- adequacy of the Smith-Geddes boundaries,
- need for acceleration type PIOs,
- importance of multiple resonances in PIO detection,
- automating the analysis process to achieve the third goal mentioned in the abstract.

An in-house Wright laboratory effort is anticipated to develop and implement a real time PIO detector based upon this work.

Acknowledgments

References

1. Bendat, Julius S. and Allan G. Piersol, *Engineering Applications of Correlations and Spectral Analysis*, John Wiley & Sons, New York, 1980.

2. Mitchell, David G. and Hoh, Roger H. , *Development of a Unified Method To Predict Tendencies for Pilot-Induced Oscillations*, WL-TR-95-3049, June 1995.

3. Klyde, David H., Duane T. McRuer, and Thomas T. Myers, *Unified Pilot-Induced Oscillation Theory, Volume 1: PIO Analysis with Linear and Non linear Effective Vehicle Characteristics, Including Rate Limiting*, WL-TR-96-3028, December 1995.

4. Preston, Jeff D., Hodgkinson, John, and Buckley, Jim, *Unified Pilot-Induced Oscillation Theory, Volume II: Pilot Induced Oscillation Criteria Applied to Several McDonnell Douglas Aircraft*, WL-TR-96-3029, December 1995.

5. Anderson *Unified Pilot-Induced Oscillation Theory, Volume 1: PIO Analysis with Linear and Non linear Effective Vehicle Characteristics, Including Rate Limiting*, WL-TR-96-3028, December 1995.

6. Bailey, Randall E. and Bidlack, Timothy, J., *Unified Pilot-Induced Oscillation Theory, Volume IV: Time Domain Neal-Smith Criteria*, WL-TR-96-3031, December 1995.

7. Bjorkman, Eileen A. *Flight Test Evaluation of Techniques to Predict Longitudinal Pilot Induced Oscillations*, M.S. Thesis, Air Force Institute of Technology, AFIT/GAE/AA/86J-1, December 1986.

8. *Military Standard, Flying Qualities of Piloted Aircraft*, MIL-STD-1797A, revised 28 June 1995.

9. Welch, Peter D., The Use of Fast Fourier Transform for Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms, *IEEE Trans. Audio and Electroacoust.*, vol. AU-15, pp. 70-73, June 1967.

10 Smith, Ralph H., "The Smith-Geddes Handling Qualities Criterion," 1994 SAE Aerospace Atlantic Conference, Dayton, OH, April 20, 1994.

Table 1. Summary of Numerical Results for Longitudinal PIOs.

	PHASE LAG ANALYSIS				
Long. PIO from Flight					
Aircraft	wr	ThetaWrdeg	qWrdeg	nzWrdeg	phi
PIO in time history					
T-38	7.49	-268	-234	-248	-355
Shuttle(t=24-30)	2.40	NaN	-153	NaN	NaN
Shuttle(t=26-32)	3.40	NaN	-171	NaN	NaN
B2 (t=25-31.7)	3.63	-165	-8	-138	-190
F8 (t=8.1-20.6)	2.79	-220	-144	-185	-225
YF-22	3.25	-213	-121	-197	-243
Bjorkman 2-5	2.57	-208	-125	-196	-232
Bjorkman 2-8	4.33	-219	-126	-174	-236
Bjorkman 3-12	1.91	-162	-72	-126	-154
Bjorkman 3-13	3.21	-184	-130	-178	-224
Bjorkman 5-9	3.51	-227	-122	-185	-235
Bjorkman 5-10	2.81	-195	-119	-176	-216
HAVE, 4/13,#21(Bj3-13)	1.32	-277	-196	-261	-279
HAVE, 4/13,#22(Bj3-13)	1.13	NaN	-195	-271	-287
HAVE, 4/13,#23(Bj3-13)	1.76	-320	-236	-322	-348
no PIO in time history					
Bjorkman 2-B	4.95	-123	-74	-137	-207
Bjorkman 2-1(wr=1.1)	1.12	-70	37	-93	-109
(wr=5.4)	5.43	-266	-97	-178	-256
Bjorkman 3-D	6.88	-180	-89	-32	-130
Bjorkman 3-1	1.78	-67	14	-67	-93
Bjorkman 3-3	3.51	-119	-47	-114	-165
Bjorkman 3-6 (wr=8.9)	8.87	-203	-108	-36	-163
(wr=3.0)	3.00	-105	-13	-50	-93
Bjorkman 4-1 (wr=6)	5.97	-174	-94	-141	-227
(wr=10)	10.00	-277	-122	-340	-483
(wr=2.5)	2.52	-120	-8	-92	-128
Bjorkman 4-2	4.83	-155	-65	-129	-198
Bjorkman 5-1	2.37	-100	-35	-100	-133
maybe PIO in time history					
Bjorkman 2-7(t=14-29)	4.29	-197	-119	-138	-200
(t=22-29, wr=4.3)	4.28	-194	-119	-144	-205
(t=22-29, wr=6.1)	6.12	-241	-149	-120	-208
Bjorkman 3-8	7.03	-246	-155	-117	-217
Bjorkman 5-11 (wr=2.5)	2.50	-202	-113	-142	-178
(wr=1.9)	1.87	-147	-56	-108	-135
(wr=3.2)	3.32	-229	-137	-145	-193

Long. PIO from Flight	Overall Results					
	Flight Ratings		FFT Classification			
Aircraft	Act. PIOR	FQ Lev	Sus_PIO	Correct		
PIO in time history						
T-38			1	y		
Shuttle(t=24-30)			1	y		
Shuttle(t=26-32)			1	y		
B2 (t=25-31.7)			1	y		
F8 (t=8.1-20.6)			1	y		
YF-22			1	y		
Bjorkman 2-5	4.33(4,4,5)	3	1	y	single dominant freq	
Bjorkman 2-8	4(4,4,4)	3	1	y	single dominant freq	
Bjorkman 3-12	4.5(4,5)	3	0	n	clear reson., not quite, miss	
Bjorkman 3-13	4.5(4,5)	3	1	y	single dominant freq	
Bjorkman 5-9	4(4,4)	3	1	y	single dominant freq	
Bjorkman 5-10	5(5,5)	3	1	y	single dominant freq	
HAVE, 4/13,#21(Bj3-13)	5		1	y		
HAVE, 4/13,#22(Bj3-13)	5		1	y		
HAVE, 4/13,#23(Bj3-13)	5		1	y		
no PIO in time history						
Bjorkman 2-B	2(3,2,2,1)	1	0	y	multiple resonances	
Bjorkman 2-1(wr=1.1)	1(1,1,1)	1	0	y	rerun wr=6	
(wr=5.4)			1	n		
Bjorkman 3-D	1(1,1)	1	0	y	multiple resonances	
Bjorkman 3-1	2.33(3,2,2)	2	0	y	multiple reson. w/dominant	
Bjorkman 3-3	1.68(3,1,1)	1	0	y	multiple resonances	
Bjorkman 3-6 (wr=8.9)	2(2,2)	2	1	n	reurn wr=3	
(wr=3.0)			0	y		
Bjorkman 4-1 (wr=6)	1(1,1,1)	1	1	n	multip. reson., freqs no match
(wr=10)			1	n		
(wr=2.5)			0	y	this is the 'best' wr	
Bjorkman 4-2	1.33(1,1,2)	1	0	y	single dominant freq	
Bjorkman 5-1	1(1,1)	2	0	y	multiple resonances	
maybe PIO in time history						
Bjorkman 2-7(t=14-29)	3(4,3,2)	2	1	y	reson. no mtch, rerun last 6s	
(t=22-29, wr=4.3)			1	y		
(t=22-29, wr=6.1)			1	y		
Bjorkman 3-8	3.68(4,3,4)	3	1	y	single dominant freq??	
Bjorkman 5-11 (wr=2.5)	3(2,4,3)	3	1	y	different but clear peaks	
(wr=1.9)			0			
(wr=3.2)			1			

Table 2. Summary of PIO Classification Results

Configurations	Results

14 PIO-prone Configs
  Bjorkmans configs with
    PIOR consistently >4
T-38, Shuttle, B-2, F-8, YF-22
3 HAVE PIOs (Bjorkman 3-12)
	correct classification: 13
	wrong classification: 1
	  (Bjorkman 3-12)
	(There was a clear single dominant
	 frequency.)

9 not PIO-prone Configs
  Bjorkman Configs with
 	
    PIOR consistently<4
	correct classification: 6
	correct by choice of wr: 3
	wrong classification: 0

3 maybe PIO Configs
  at least 1 PIOR >=4
  (Bjorkman 2-7,
    Bjorkman 3-8,
    Bjorkman 5-11)
	correct by choice of wr: 3
	wrong classification: 0
	Config(PIOR)
	Bj2-7 (4,3,2) PIO-prone at 3/3 wrs
	Bj3-8 (4,3,4) PIO-prone at 1/1 wrs
	Bj5-11(2,4,3) PIO-prone at 2/3 wrs
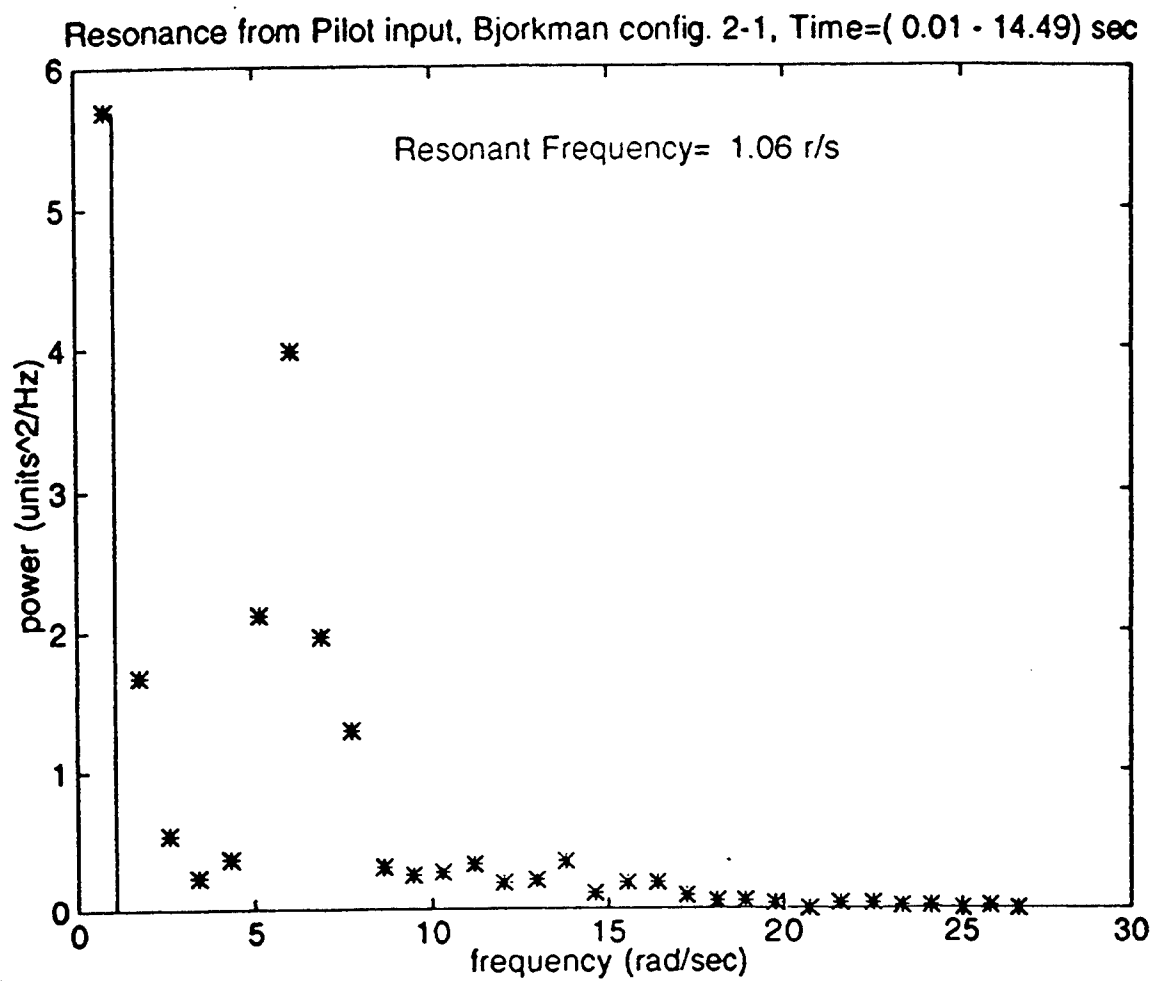
Figure 1. Power spectrum showing multiple resonances and low frequency power
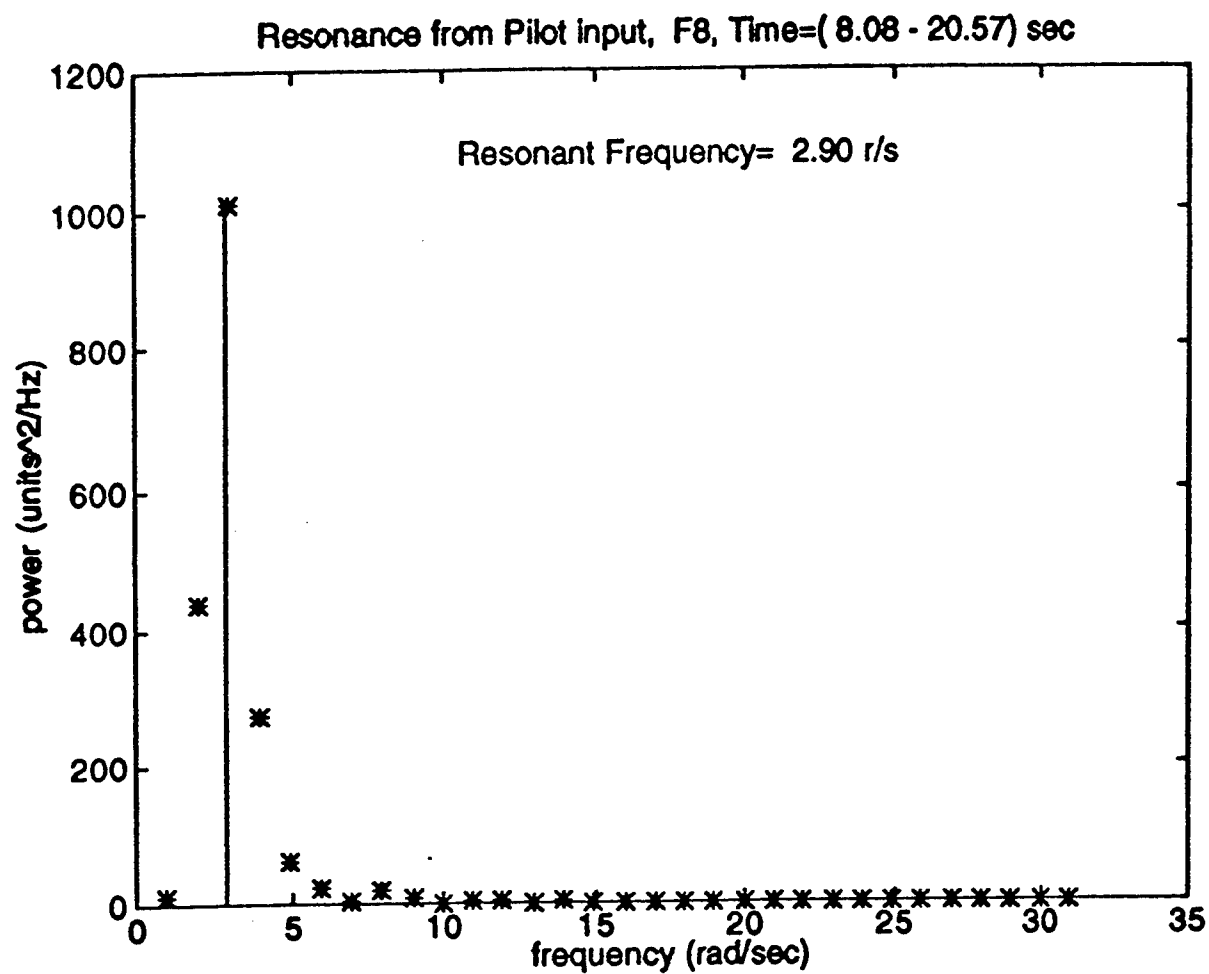
Figure 2. Power spectrum showing a single dominant frequency
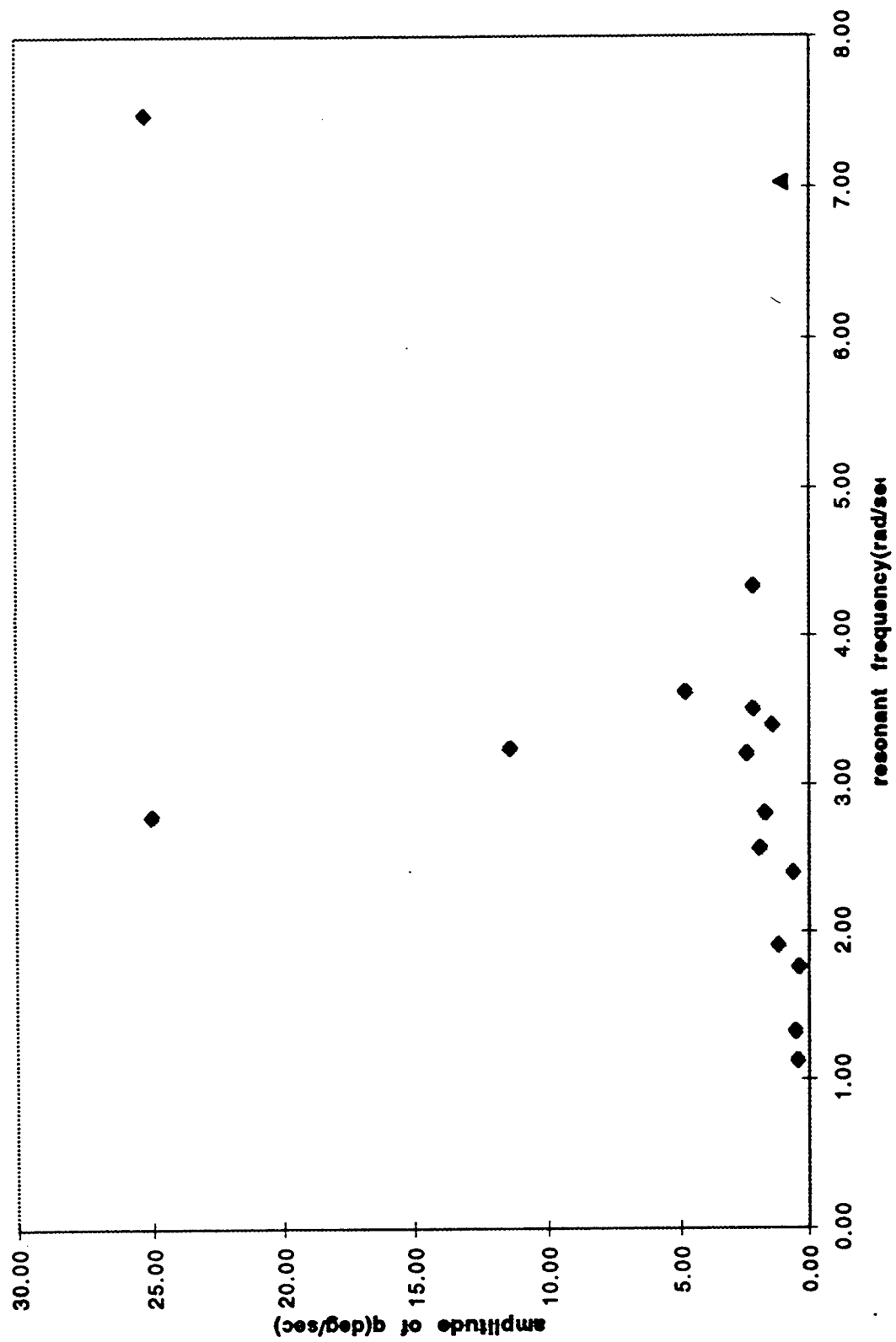
Figure 3. Properties of Longitudinal PIOs
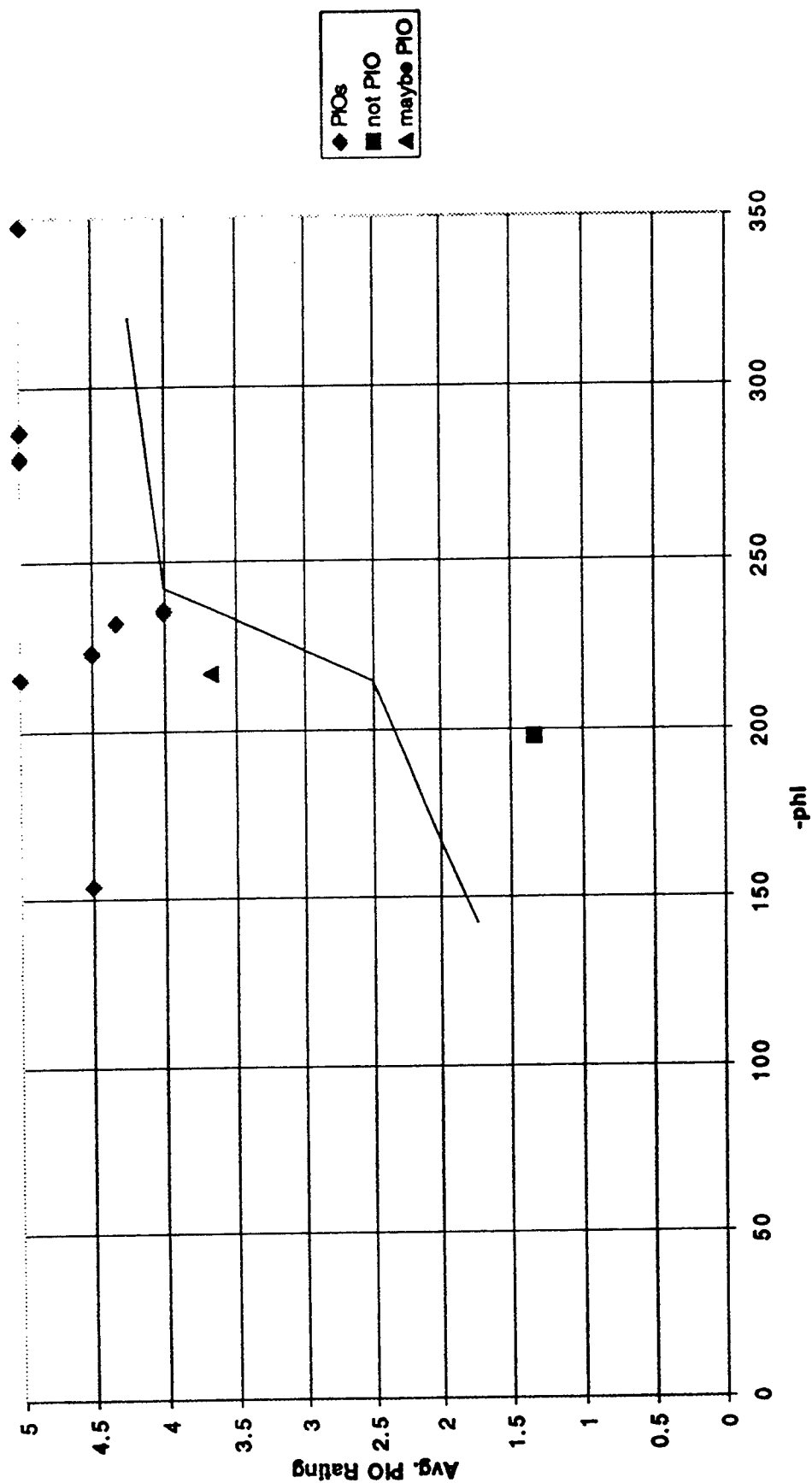
Figure 4. Actual average PIO ratings versus phi for Bjorkman records
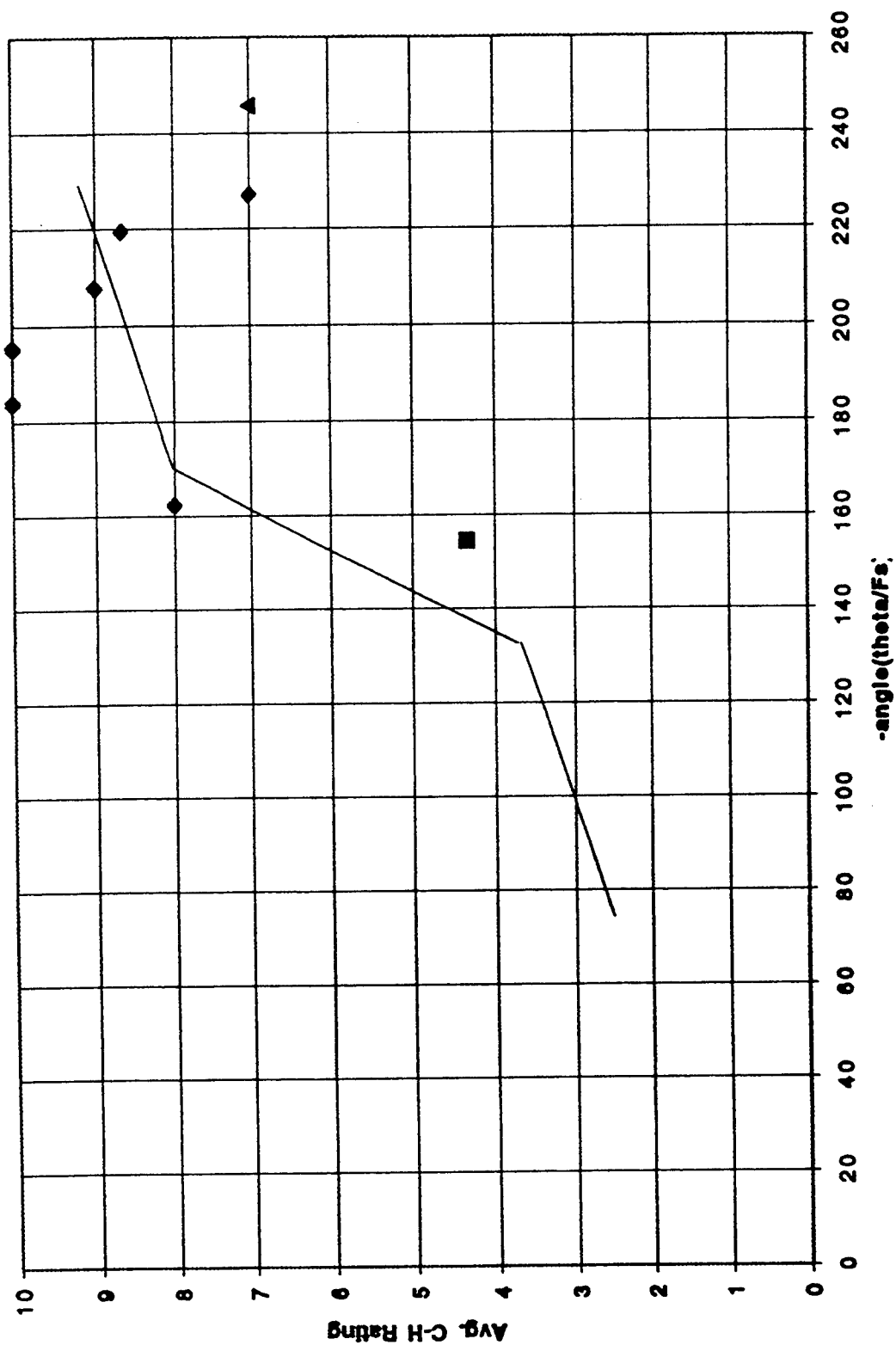
Figure 5. Actual average Cooper-Harper ratings versus phase of $\Theta(j\omega_r)/Fs(j\omega_r)$ for Bjorkman records
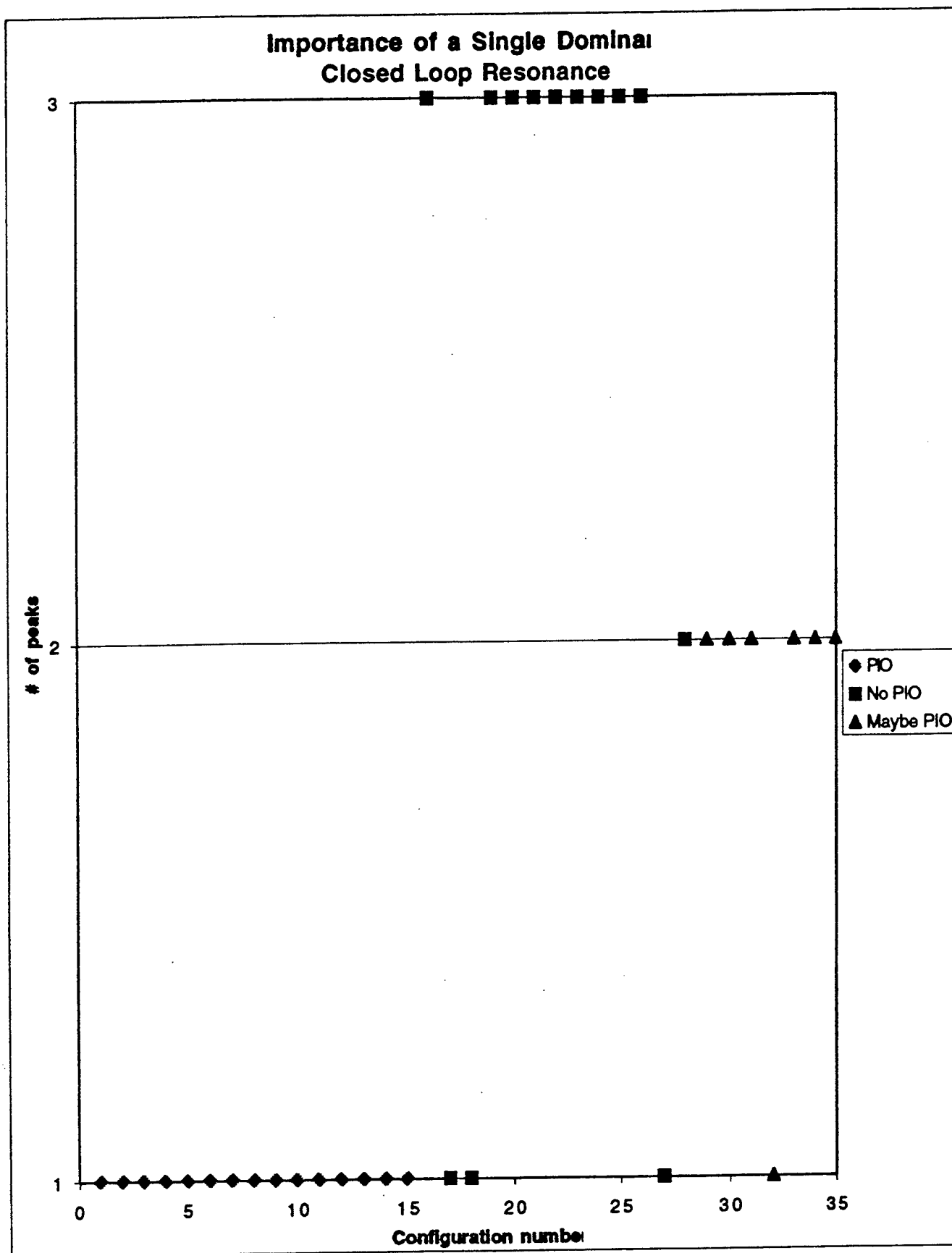
Figure 6. The importance of a single dominant frequency for PIO
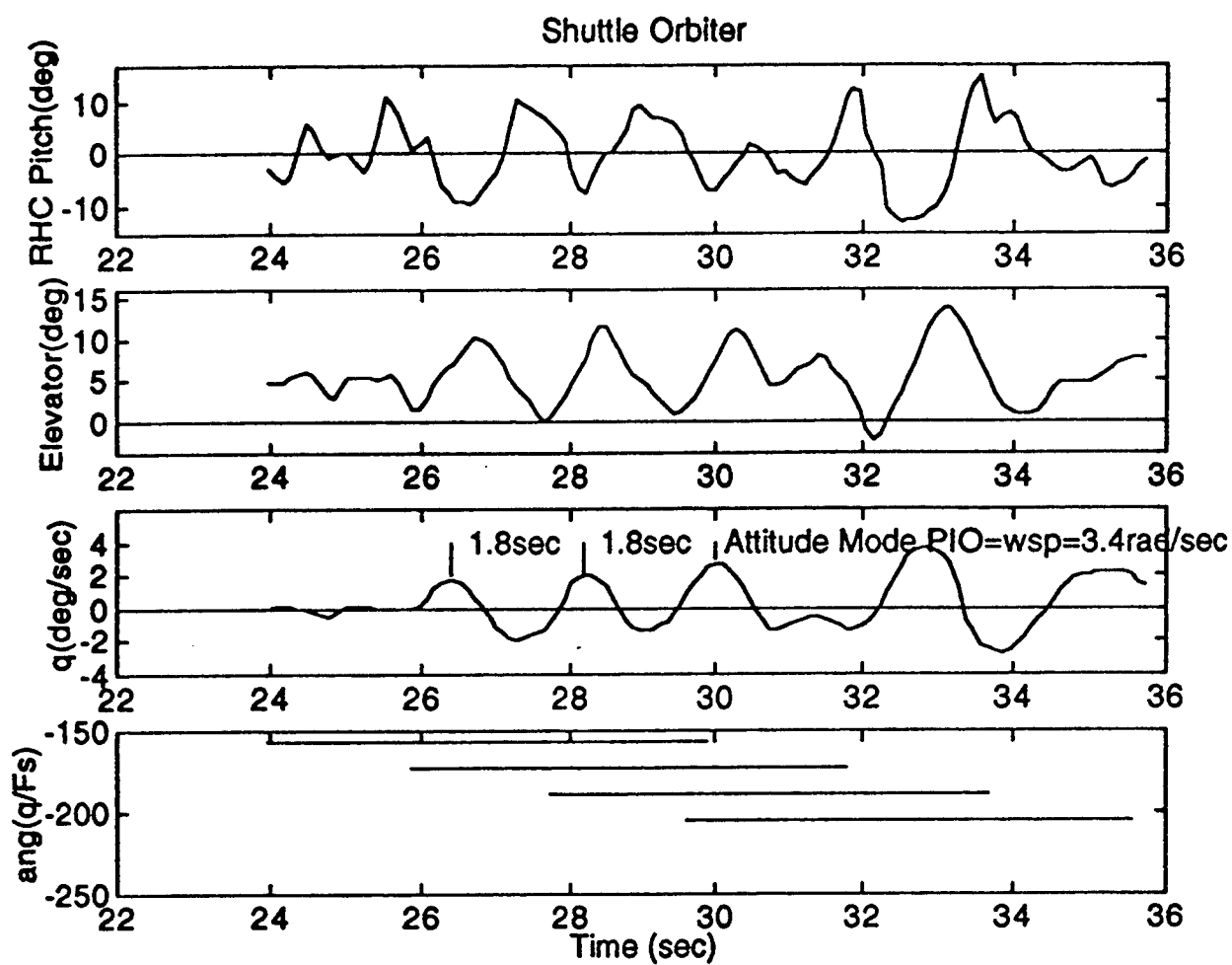
Figure 7. Shuttle record showing significant time variation of phase angle of $q(j\omega_r)/F_s(j\omega_r)$
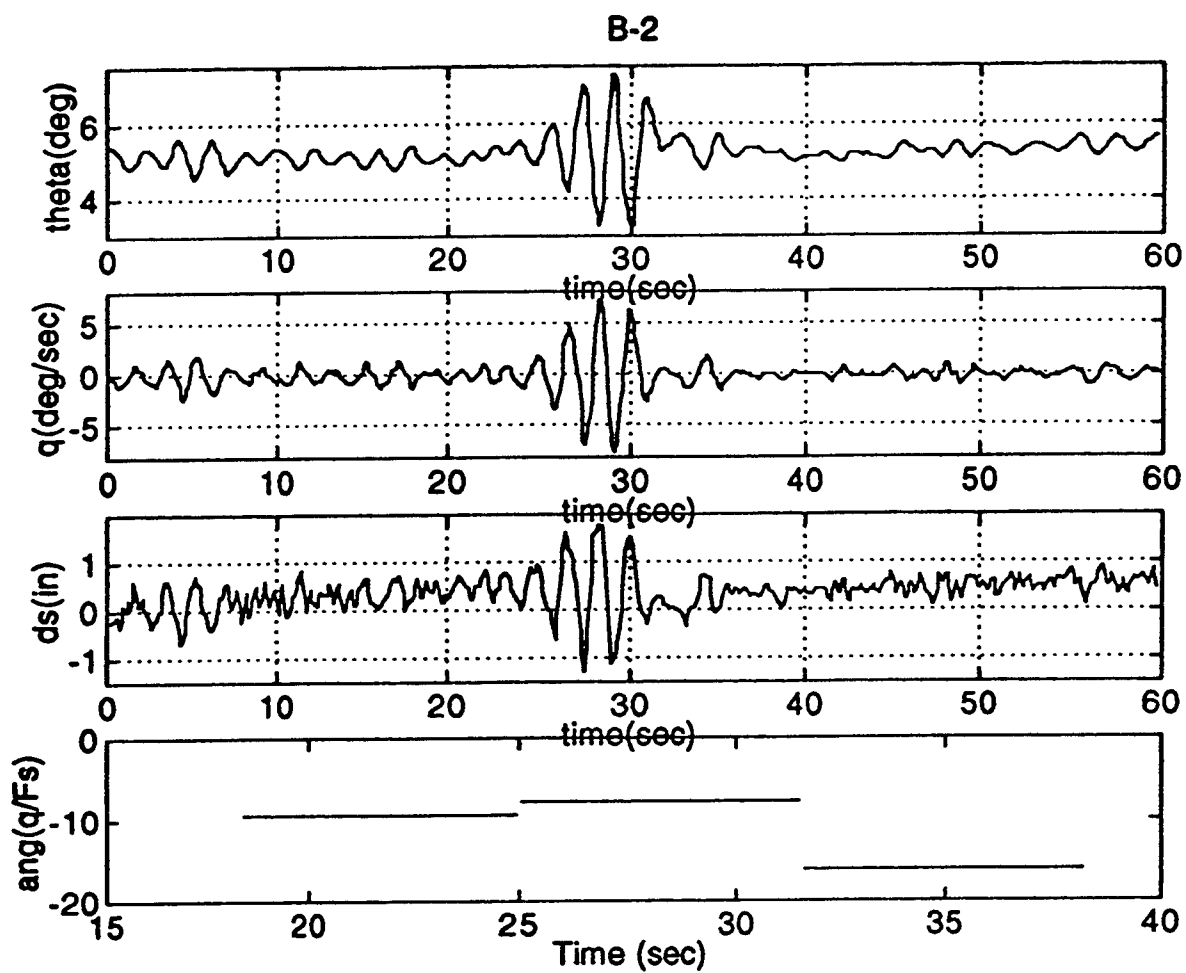
Figure 8. B2 record showing insignificant time variation of phase angle
of $q(j\omega_r)/F_s(j\omega_r)$

SYNTHESIS OF NOVEL THIRD ORDER NONLINEAR OPTICAL MATERIALS

Lawrence L. Brott

Graduate Student

Department of Materials Science and Engineering

497 Rhodes Hall

University of Cincinnati

OH 45221-0012

Final Report for:

Graduate Student Research Program

WL/MLBP

Wright Patterson AFB

September 1996

SYNTHESIS OF NOVEL THIRD ORDER NONLINEAR OPTICAL MATERIALS

Lawrence L. Brott

Department of Materials Science and Engineering

University of Cincinnati

ABSTRACT

Synthesis of third order nonlinear optical (NLO) polymers represents an exciting field with the resulting chromophore containing materials being used for two-photon-pumped upconverted lasing or as optical limiters. In this research, three novel NLO chromophores are developed by incorporating a fluorene molecule in their backbone with either thiophene and pyridine end groups which act as electron donating or withdrawing groups respectively. Long alkyl chains are attached to the C-9 carbon on the fluorene backbone to aid in the chromophore's solubility in the host polymer.

SYNTHESIS OF NOVEL THIRD ORDER NONLINEAR OPTICAL MATERIALS

Lawrence L. Brott

Department of Materials Science and Engineering

University of Cincinnati

INTRODUCTION

Interest in third order nonlinear optical (NLO) polymers has increased dramatically in the past few years as potential applications (for example, optical memory or cancer therapy) begin to be realized [1,2]. Third order NLO compounds typically, though not necessarily, are symmetrical molecules comprised of multiple conjugated bonds. Present polymers however still lack the nonlinearity and stability that are necessary for them to be used commercially. The objective of this research is to design and synthesize three new fluorene-containing compounds (**1**-**3**) for the use by the U.S. Air Force that will optimize nonlinearity while remaining thermally and structurally stable. Compound **1** is constructed by using an alkylated fluorene group to bridge two electron-donating thiophene rings, while compound **2** uses two electron-withdrawing pyridine groups. Compound **3** is similar to compound **2** except for the additional conjugated bonds.
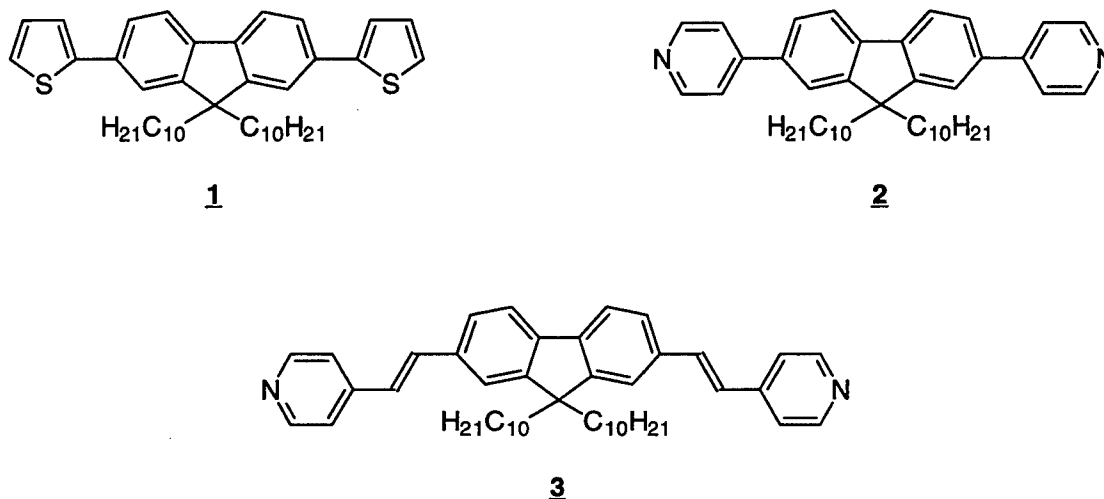
Fig. 1   Fluorene-containing chromophores.

EXPERIMENTAL DETAILS

PREPARATION OF THE FLUORENE BACKBONE

The central building block of the three chromophores (1-3) is the fluorene molecule. In order to make it more soluble and easier to handle, two long alkyl chains were added to the C-9 position of the fluorene. This approach allowed the C-2 and C-7 carbons to still be sterically unhindered and reactive. As seen in figure 2, fluorene was first treated with BuLi, complexed with TMEDA [3,4], and then treated with bromodecane to obtain monoalkylated fluorene. The monoalkylated fluorene, without isolation and purification, was further reacted with a second equivalent of BuLi and bromodecane to obtain dialkyl compound 5. The product was purified by column chromatography followed by distillation under reduced pressure to remove any residual bromodecane.
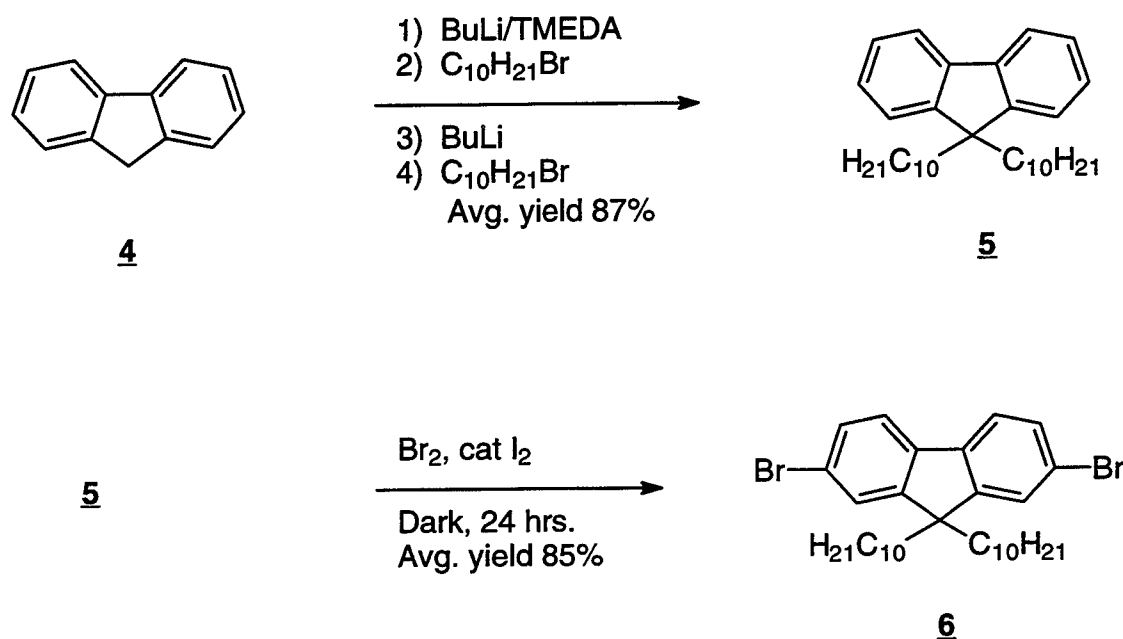
Fig. 2   2,7-dibromo-9,9-di-n-decylfluorene 6 synthesis.

The next key step was to synthesize 2,7-dibromofluorene 6 [5]. Compound 5 was reacted with bromine in the presence of a small amount of iodine and in the complete absence of light. The dibromofluorene 6 was then purified by column chromatography.

ARYL COUPLING

The thiophene and pyridine end groups in compounds **1** and **2** were attached to the dibromofluorene **6** by a reaction as shown in figure 3 [6,7]. In this process, a (tributylstannyl)aryl **7** is reacted with the dibromofluorene **6** in the presence of a palladium catalyst to produce a coupled product.

2 Ar—Sn(Bu)₃ + Br—[fluorene, H₂₁C₁₀ C₁₀H₂₁]—Br $\xrightarrow[\substack{\text{Toluene} \\ 111°C}]{\text{Pd(PPh}_3)_4}$ Ar—[fluorene, H₂₁C₁₀ C₁₀H₂₁]—Ar

7   **6**

Fig. 3   Carbon-carbon bond formation.
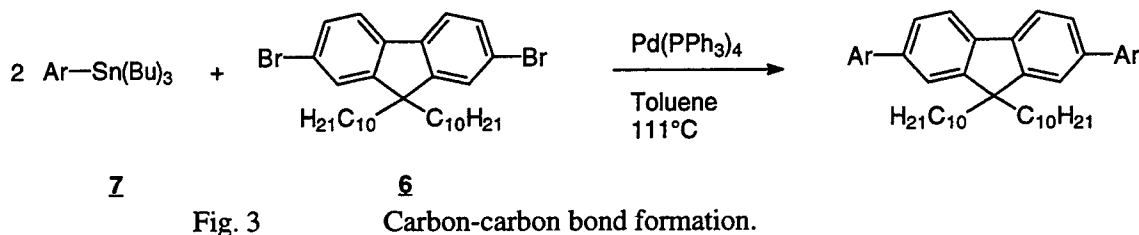
2-(TRIBUTYLSTANNYL)THIOPHENE

Chromophores **1** and **2** incorporate one of two different end groups that act as either electron donors and acceptors. In this research, a thiophene ring will act as the donor while the pyridine ring will be the acceptor. Consequently, two additional products needed to be synthesized. Figure 4 describes the preparation of the thiophene compound.

[thiophene]—Li + (Bu)₃SnCl $\xrightarrow[\substack{25°C \\ 47\% \text{ yield}}]{\substack{\text{THF} \\ 36\text{ hours}}}$ [thiophene]—Sn(Bu)₃

8
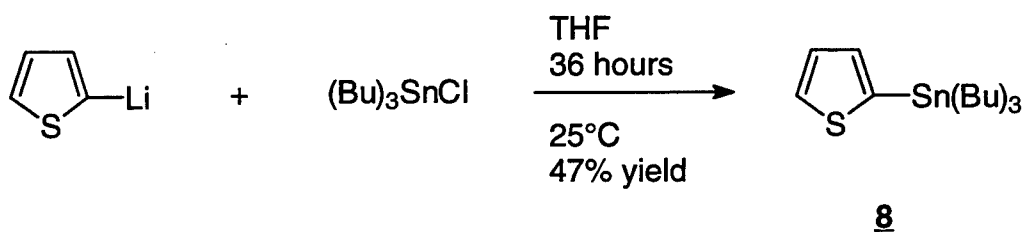
Fig. 4   Stannyl-thiophene **8** synthesis.

In order to remove any unreacted tin compound, the THF was rotavaped off, and then hexanes added. An aqueous potassium fluoride solution was added to the flask and stirred vigorously for an hour. The layers were separated with the organic layer being dried with $MgSO_4$, filtered, and the product was then purified by column chromatography.

4-(TRIBUTYLSTANNYL)PYRIDINE

The pyridine stannation is similar to the stannyl-thiophene preparation, except that the 4-bromopyridine hydrochloride must be neutralized and lithiated before the tributyltin chloride can be added. Both of these objectives can be achieved by adding two equivalents of butyl lithium as described in figure 5.
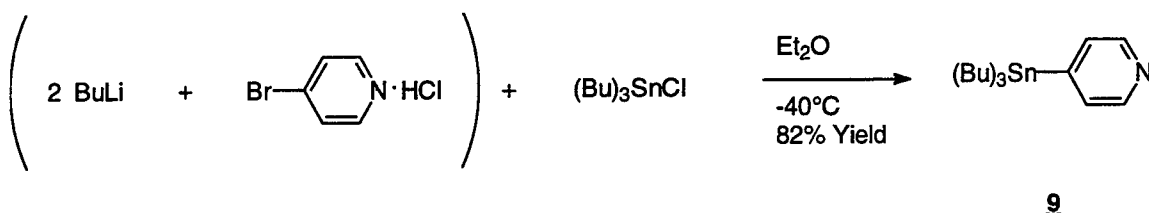
Fig. 5   Stannyl-pyridine **9** synthesis.

A slurry of the pyridine salt was vigorously stirred in dry diethyl ether which was cooled to -40°C. Butyl lithium was slowly dripped in so as to not raise the internal temperature. The solution was then allowed to stir an additional twenty minutes before the tributyltin was added. The remaining work-up and purification was similar to the stannyl-thiophene **8**.

CHROMOPHORE SYNTHESES
2,7-(2-THIENYL)-9,9-DI-*n*-DECYLFLUORENE

The chromophore **1** was completed by coupling two equivalents of the stannyl-thiophene **8** with the dibromofluorene **6** as shown in figure 6 [8].
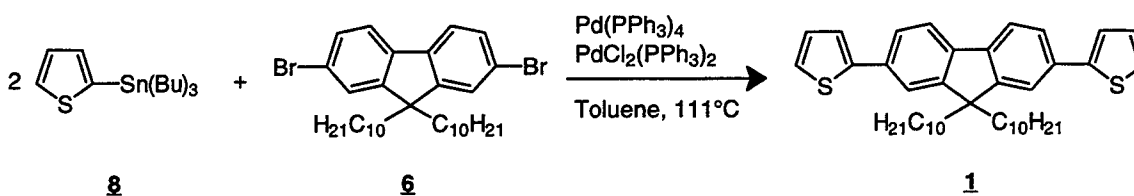
Fig. 6   Compound **1** synthesis.

After adding compounds **8** and **6** together in toluene, palladium catalyst was added and the solution was refluxed for four hours. It was determined that the reaction was completed when the solution turns black. The work-up was the same as the preparation of the stannyl-thiophene

8, and the product was purified by column chromatography to obtain a white, waxy solid **1** (mp=59.9-60.8°C) in 71% yield.

2,7-(4-PYRIDYL)-9,9-DI-*n*-DECYLFLUORENE

Compound **2** was synthesized in a similar manner as compound **1**, except that triphenylphosphine was also used in addition to the two palladium catalysts (see figure 7).

Fig. 7   Compound **2** synthesis.
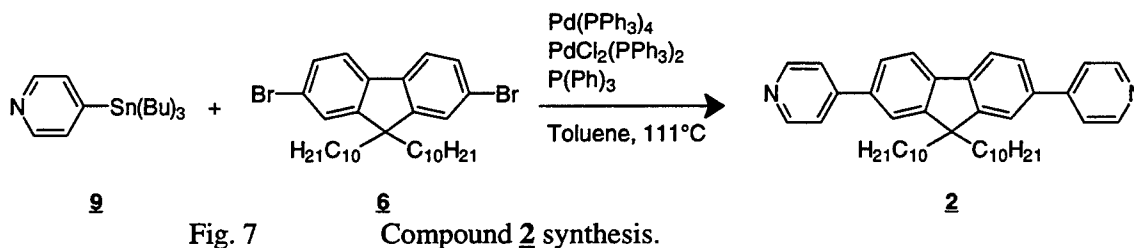
After adding compounds **9** and **6** together in toluene, two different palladium catalysts with triphenylphosphine were added and the solution was refluxed overnight. As before, the reaction was complete when the solution turned black. The work-up was the same as the preparation of the stannyl-thiophene **8**, and the product was purified by column chromatography to obtain a light brown oil in 19% yield.

4'-(9,9-DI-*n*-DECYLFLUORENE-2,7-FLUORENEDIYLDI-2,1-ETHENEDIYL)-BIS-PYRIDINE

Unlike the previous syntheses, compound **3** was synthesized using a technique known as the Heck reaction.

Fig. 8   Compound **3** synthesis.
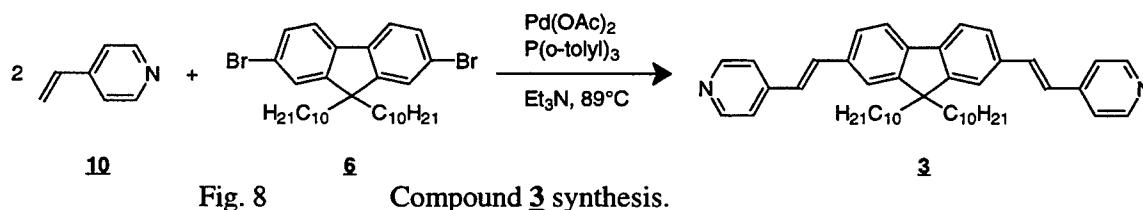
After combining the starting materials, the solution was allowed to reflux overnight. As before, the reaction was complete when the solution turned black. The work-up included removing the

4-7

triethylamine solvent and replacing it with methylene chloride. The organic solution was then washed twice with water, separated, dried with $MgSO_4$, filtered, and then concentrated. The product was purified on an alumina column to yield a light yellow oil in 72% yield.

CONCLUDING REMARKS

The synthetic methods to yield three novel third order NLO material containing fluorene groups has been successfully developed. Under a University of Cincinnati / WL/WPAFB educational partnership agreement, the synthesis, purification and characterization of these compounds will be carried out. The author will continue this successful investigation on NLO materials both at the University and WL/WPAFB under this joint agreement and third order NLO measurements of this material will be carried out in the coming months.

ACKNOWLEDGMENTS

It is a pleasure to thank the individuals who have made the research visit to WL/MLBP both a fruitful and pleasurable experience. In particular, thanks go to Mr. Bruce Reinhardt, Dr. Bob Evers, Dr. Ted Helminiak, Ms. Lisa Denny, Ms. Marilyn Unroe, Ms. Ann G. Dillard, and Dr. Ram Kannan for all their help and kind hospitality.

BIBLIOGRAPHY

1. Prasad, P.N. and Williams, D.J., *Introduction to Nonlinear Optical Effects in Molecules and Polymers*, Wiley-Interscience, New York, 1991.

2. Reinhardt, B.A., *Trends in Polymer Science*, **1996**, 4(9), 287.

3. Yasuda, H., Walczak, M., Rhine, W. and Stucky, G., *J. Organomett. Chem.*, **1975**, 90, 123-31.

4. Zerger, R., Rhine, W. and Stucky, G., *J. Am. Chem. Soc.*, **1974**, 96, 5441-8.

5. Rehahn, M., Schluter, A.-D. and Feast, J.W., *Synthesis*, **1988**, 386-389.

6. Takahashi, K. and Nihira, T., *Bull. Chem. Soc. Jpn.*, **1992**, 65, 1855-1859.

7. Yamamoto, Y. and Yanagi, A., *Heterocycles*, **1981**, 16(7), 1161-1164.

8. Stille, J.K., McKean, d.R., Parrinell, G. and Renaldo, A.F., *Journal of Organic Chemistry*, **1987**, 52, 422-424.

PROBING THE UNIQUE PROPERTIES OF A SUPERCRITICAL FLUID: FLUORESCENCE RADIATIVE RATE CONSTANT DEPENDENCE ON SUPERCRITICAL FLUID DENSITY, AND BIMOLECULAR QUENCHING EFFICIENCIES AS OBSERVED BY FLUORESCENCE LIFETIME MEASUREMENTS IN SUPERCRITICAL CARBON DIOXIDE

Christopher E. Bunker
Graduate Student
Department of Chemistry

Clemson University
Hunter Hall
Clemson, SC 29634-1905

PROBING THE UNIQUE PROPERTIES OF A SUPERCRITICAL FLUID: FLUORESCENCE RADIATIVE RATE CONSTANT DEPENDENCE ON SUPERCRITICAL FLUID DENSITY, AND BIMOLECULAR QUENCHING EFFICIENCIES AS OBSERVED BY FLUORESCENCE LIFETIME MEASUREMENTS IN SUPERCRITICAL CARBON DIOXIDE

Christopher E. Bunker
Graduate Student
Department of Chemistry
Clemson University

Abstract

1. The dependence of the fluorescence radiative rate constant k_F of the model fluorophors 9-cyanoanthracene, 9,10-diphenylanthracene, and anthracene on the reduced density and refractive index of supercritical carbon dioxide at 35 °C is investigated. Fluorescence lifetimes are determined using the time-correlated single photon counting method. The k_F results support previous studies demonstrating anomalous behavior in the low density regions of a supercritical fluid, and indicate a strong system dependence. 9-cyanoanthracene and 9,10-diphenylanthracene both display an unusual behavior of k_F in the low density region of supercritical carbon dioxide, while anthracene results agree well with radiative rate theory. 2. The quenching of 9-cyanoanthracene, 9,10-diphenylanthracene, anthracene, and perylene fluorescence by carbon tetrabromide is investigated in supercritical carbon dioxide as a function of density. Fluorescence lifetimes of the model fluorophors at different carbon tetrabromide concentrations are determined. The fluorescence decays adhere closely to single-exponential functions and the carbon tetrabromide concentration dependence follows the Stern-Volmer relationship. The quenching of the fluorescence lifetime for 9-cyanoanthracene and 9,10-diphenylanthracene is unusually efficient at near-critical densities. The results are in support of a solute-solute clustering mechanism in which the enhance bimolecular reactions are attributed to a higher local quencher concentration then the bulk under near-critical solvent conditions. At the same time, the quenching of anthracene and perylene display no unusual behavior and are apperently diffusion controlled throughout the density region examined.

PROBING THE UNIQUE PROPERTIES OF A SUPERCRITICAL FLUID: FLUORESCENCE RADIATIVE RATE CONSTANT DEPENDENCE ON SUPERCRITICAL FLUID DENSITY, AND BIMOLECULAR QUENCHING EFFICIENCIES AS OBSERVED BY FLUORESCENCE LIFETIME MEASUREMENTS IN SUPERCRITICAL CARBON DIOXIDE

Christopher E. Bunker

Introduction

Supercritical fluids are solvent systems that have received a great deal of scientific interest.[1-3] The most important properties of a supercritical fluid are the low densities between those of a gas and a liquid and which are easily tunable with changes in pressure or temperature, and the local density effects. The low densities and tunability of supercritical fluids make them ideal for use as reaction media.[4] Greater reaction efficiencies due to the improved mass transport properties and better product selectivities and control are anticipated. It is in the area of local density effects that the greatest interest in supercritical fluids has been observed. Local density effects can be described by two phenomena. The first, solvent-solute clustering, describes a situation in which the local density of solvent molecules about a solute is greater than the bulk. Experimental measurements of solvent clustering show density enhancements of as much as 2-3 times over those of the bulk density in the near-critical region.[5-20] The second, solute-solute clustering, refers to a situation in which the local concentration of solute molecules is greater than the bulk concentration. Experimental results from studies of solute-solute clustering have been inconsistent. Solute-solute clustering was initially proposed based on the observed entrainer effects[21] and on unusually efficient excimer and exciplex formation in supercritical fluids.[6a,22] Further studies supported the observation of enhanced solute-solute interactions in supercritical fluids.[23,24] Recently, a report of the bi-molecular quenching reaction of 9,10-bis(phenylethynyl)anthracene (BPEA) with carbon tetrabromide (CBr_4) in supercritical carbon dioxide demonstrated concentration enhancements of the quencher of ~ 100% in the near-critical density region.[25] There have also been a fair number of experimental studies that have shown no apparent enhancements in solute-solute interaction above those predicted by theory.[26-29] Obviously, a better understanding of the nature of a supercritical fluid in the near-critical density region and its effects on chemical reactions is required. In this study we have conducted experiments designed to probe the unique properties of a supercritical fluid.

Specifically, we have studied the dependence of the fluorescence radiative rate constant of several model fluorophors on the refractive index of supercritical carbon dioxide and the bimolecular quenching efficiencies of several anthracene derivatives by carbon tetrabromide in supercritical carbon dioxide.

Experimental

Materials. 9-Cyanoanthracene, 9,10-diphenylanthracene, anthracene, and perylene (Aldrich, 99%+) were used as received. Carbon dioxide (Air Products, 99.9999%) was checked for oxygen content by comparing fluorescence quantum yields of pyrene before and after a rigorous deoxygenation procedure in which the carbon dioxide was repeatedly passed through oxygen traps (Alltech Associates, Inc.). Because the results indicated no oxygen quenching, the carbon dioxide was used without further purification.

Measurements. Fluorescence decays were measured using a time-correlated single photon counting apparatus. The excitation source was a Spectra Physics Tsunami titanium-sapphire mode locked laser, pulse picked at 4 mHz and doubled with excitation wavelengths tuned between 360 to 400 nm. The detector consisted of a Hamamatsu microchannel plate in a thermoelectrically cooled housing. The time resolution of the setup as measured by the instrument response function was ~ 100 ps. Fluorescence decays were detected with use of a monochrometer for wavelength selection. Large emission slits (2 mm) were employed for improved light gathering ability. Fluorescence lifetimes were determined from observed decay curves and instrument response functions using a standard deconvolution method.

The high-pressure setup has been described in detail.[26] Briefly, the optical cell consisted of a cubic shaped stainless steel block with four openings. Three of the openings contained quartz or sapphire windows and the fourth was for cell cleaning. Temperatures were controlled with a Omega 4200A RTD controller and two heater cartridges inserted into the cell. Pressures were monitored by a Heise-901A pressure gauge. CO_2 was added to the cell by use of a syringe pump. Samples were loaded from stock solutions of known concentration in hexane or as solid material. Excess solvent was removed with a gentile stream of nitrogen gas. The cell was then sealed and repeatedly purged with low pressure CO_2. Once equilibrated at the desired temperature and pressure, fluorescence decays could be recorded.

Results and Discussion

Fluorescence Radiative Rate Constant Dependence on SCF Refractive Index.

Fluorescence radiative rate theory predicts that the fluorescence radiative rate constant k_F will depend on solvent refractive index by a power of two,

$$k_F = n^2 * k_F^{\circ} \tag{1}$$

where n is the refractive index of the solvent and k_F° is the natural fluorescence radiative rate constant determined in a vacuum. Experimental studies in room-temperature solution generally support the n^2 dependence.[30] However, it is not always clear that one is separating the refractive index dependent behavior from changes due to the difference in the chemical environments. Supercritical fluids with their tunable solvent properties offer an ideal medium in which to study the dependence of the fluorescence radiative rate constant on refractive index. Unlike normal room-temperature solvents, the refractive index of a supercritical fluid is dependent on the fluid density. We can therefore easily change the refractive index in a supercritical fluid with a slight change in pressure. The nature of the chemical environment, however, will remain the same. Experimentally, k_F can be calculated from accurate measurements of fluorescence lifetimes and quantum yields,

$$k_F = \Phi_F / \tau_F \tag{2}$$

9-cyanoanthracene (9CA) is a fluorescent molecule with a quantum yield in room-temperature solvents of 1.0.[31] Studies in molecular beams have shown that the fluorescence quantum yield is also equal to 1.0.[31] Previously, studies of the fluorescence lifetime as a function of solvent density in supercritical ethane[32] showed an anomalous behavior of the calculated fluorescence radiative rate constant. The radiative rate constant deviated sharply upward from the predicted n^2 dependence in the low density region of supercritical ethane. While the comparison of k_F to n was made with the assumption that the fluorescence quantum yield was equal to 1.0 for supercritical ethane, it was stated that changes in the quantum yield
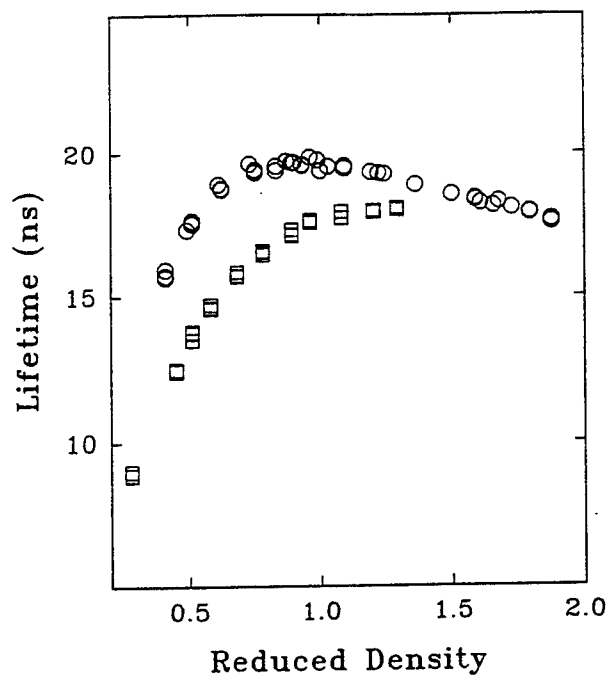
Figure 1. Fluorescence lifetimes of 9-cyanoanthracene in supercritical CO_2 at 35 °C (o) and 60 °C (□).
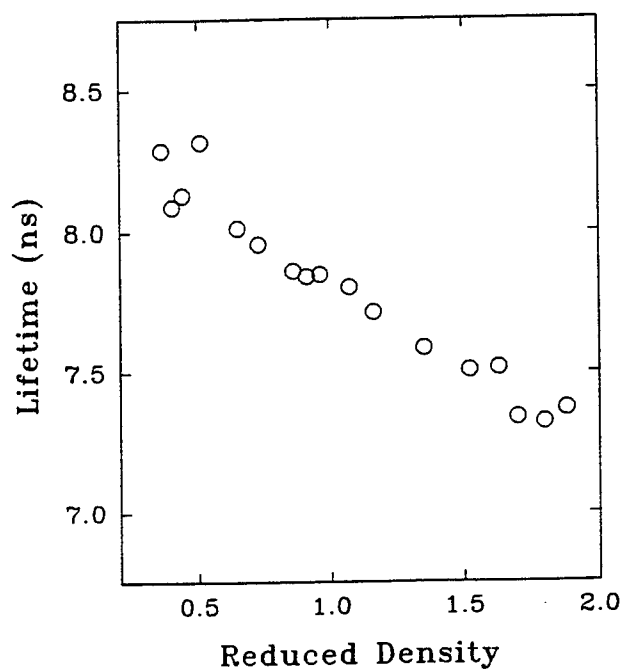
Figure 2. Fluorescence lifetimes of anthracene in supercritical CO_2 at 35 °C.

could account for the anomalous behavior of k_F. Recently, a further study of 9CA in three supercritical fluids confirmed the anomalous behavior and did make an attempt to examine the quantum yield dependence on fluid density.[33] The results were somewhat inconclusive. We have measured the fluorescence lifetimes of 9CA as a function of CO_2 reduced density at 35 and 60 °C. In addition, we have studied the fluorescence lifetime dependence of two related fluorophors, anthracene and 9,10-diphenylanthracene (DPA), on CO_2 reduced density at 35 °C. These molecules differ from 9CA in that their fluorescence quantum yields are less than 1.0. With quantum yields less than 1.0, other processes, such as intersystem crossing, can potentially play a significant role in the observed fluorescence lifetime behaviors.

Experimentally determined lifetimes for the three molecules as a function of CO_2 reduced density are shown in figures 1-3. As can be seen in figure 1, the lifetimes of 9CA increase with decreasing CO_2 ρ_r for the density region of 1.4 to 1.9. Below ρ_r of 1.4, the lifetimes roll over and decrease sharply with decreasing density. The observed behavior is in good agreement with that observed in supercritical ethane.[32] At 60 °C, the behavior is similar to that a 35 °C, only across the entire density region the lifetimes are somewhat shorter. In addition, the roll over region is somewhat broader and shifted to higher densities. If we assume that the Φ_F is equal to 1.0, although this may not be correct, we can examine the calculated radiative rate constant as a function of ρ_r and refractive index. With k_F approximated by $1/\tau_F$ and taking the natural log of equation 1 we obtain,

$$\ln(1/\tau_F) \quad = \quad x * \ln(n) \quad + \quad \ln(k_F^o) \tag{3}$$

where x is ideally equal to two. Shown in figure 4 is a plot of $\ln(1/\tau_F)$ vs $\ln(n)$ for 9CA in supercritical CO_2 at 35 and 60 °C. For the high density region (ρ_r of ~ 1.4 - 1.9), the plot is linear with a slope of 1.5. This value for the refractive index dependence is somewhat lower then predicted, however, it is in good agreement with previous results for 9CA.[31,32] Below a reduced density 1.4, the calculated radiative rate constant makes a sharp deviation upward away from the line established in the high density region. The values increase sharply with decreasing refractive index. At 60 °C, our data does not extend to high enough CO_2 densities to observe the linear portion of the plot. The low density data, however, mimics the upward deviation of
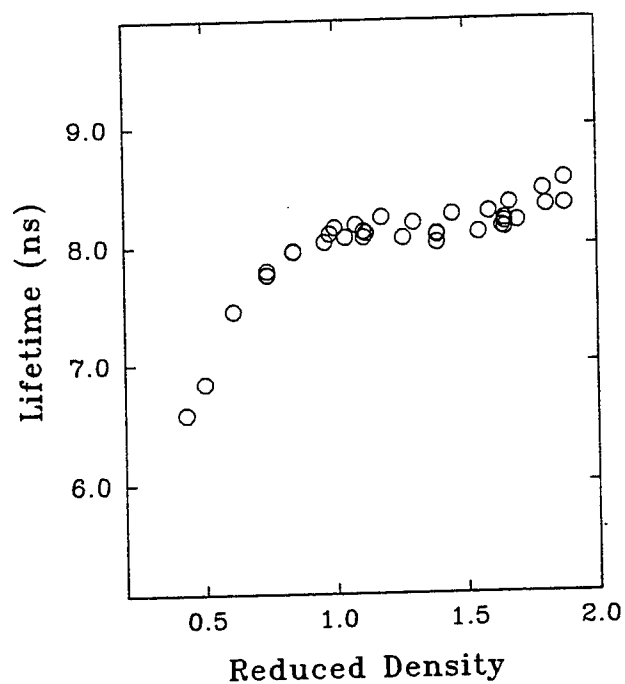
Figure 3.   Fluorescence lifetimes of 9,10-diphenylanthracene in supercritical CO_2 at 35 °C.
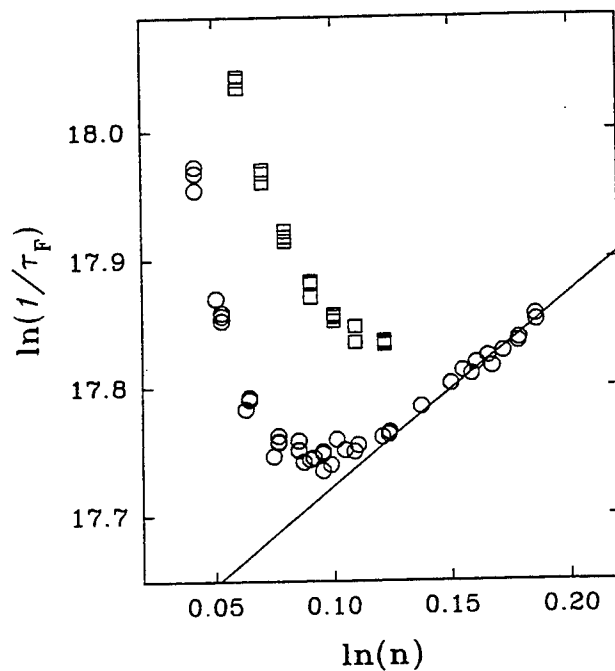
Figure 4.   A plot of $\ln(1/\tau_F)$ plotted against $\ln(n)$ for 9-cyanoanthracene in supercritical CO_2 at 35 °C (o) and 60 °C (□). The line is from a linear least squares regression.

5-8

k_F observed at 35 °C.

In order for k_F to be n^x dependent across the entire reduced density region examined, the fluorescence quantum yield must also decrease with decreasing ρ_r in the low density region in a similar manor as τ_F. It has been shown[33] that this most likely does occur, however, the results were far from conclusive. Accurate measurements of Φ_F in the low density regions of a supercritical fluid are extremely difficult. Solubilities in these regions tend to be low and fairly sensitive to changes in density. Therefore a truly quantitative determination of k_F for 9CA has yet to be completed. Assuming the quantum yield does indeed change such the k_F maintains the n^x dependence, the question becomes why do Φ_F and τ_F decrease so dramatically in the low density region? Considering the low solubilities in the low density regions, it could be possible that a self-quenching process due to reduced solubilities becomes active. As CO_2 density becomes lower, the reduced solubility could result in the formation of 9CA microcrystals or aggregates, resulting in quenching of Φ_F and τ_F. It could be expected, however, that at higher temperatures, solubilities would increase. This should cause the roll over region to shift to lower densities, not higher as is observed. Another possibility might be due to inter-system crossing to the excited triplet state. If in the low density region of the supercritical fluid, the k_{ISC} rate constant was sufficiently enhanced so as to compete for the excited singlet state population of 9CA, then the Φ_F and τ_F would indeed decrease with decreasing CO_2 density. The implication is, however, that at very low densities the Φ_F and τ_F must rapidly increase to the values determined in a vacuum.[31] The k_{ISC} rate constant would therefore be competitive in only a small density window, and as demonstrated by the literature, be independent of supercritical solvent.[33] This would be a very unusual phenomenon if found to be true. Obviously, further study is warranted.

The spectroscopic properties of anthracene in room-temperature solutions have been well studied. The deactivation of the excited singlet state of anthracene includes a significant non-radiative decay path which is primarily through inter-system crossing. The quantum yield and lifetime in non-polar hydrocarbon solvents are ~ 0.31 and 5.3 ns, respectively.[31] Anthracene is known to have a high inter-system crossing yield, approximately 0.7. With such a high inter-system crossing yield, anthracene makes an ideal probe with which to examine any effects the supercritical fluid might have on the non-radiative decay paths.
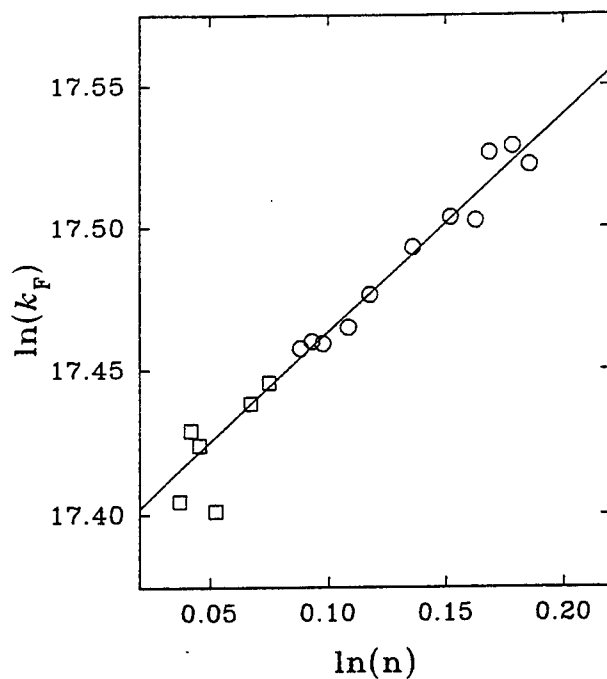
Figure 5.   A plot of $\ln(k_F)$ vs $\ln(n)$ for anthracene in supercritical CO_2 at 35 °C.   (o) True k_F values with known Φ_F's.   (□) Estimated k_F with Φ_F assumed $= 0.3$.
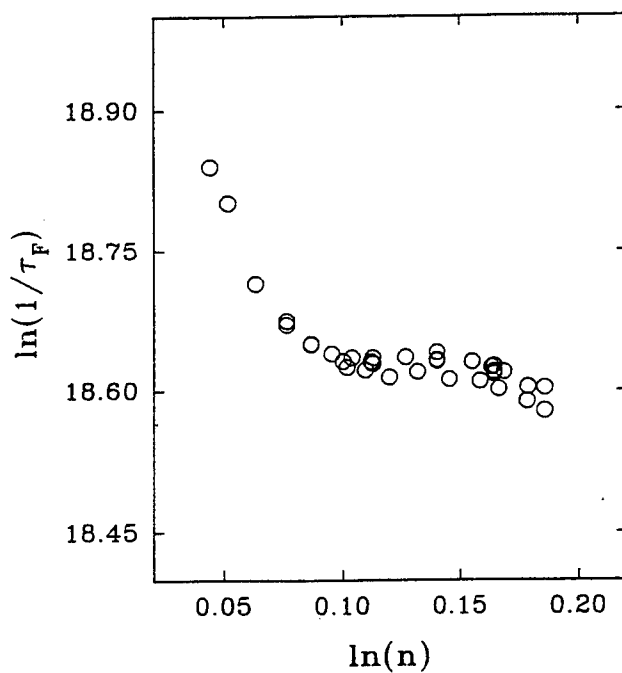
Figure 6.   A plot of $\ln(1/\tau_F)$ plotted against $\ln(n)$ for 9,10-diphenylanthracene in supercritical CO_2 at 35 °C.

Quantum yields measured as a function of CO_2 reduced density from ~ 0.8 to 1.9 at 35 °C were constant with a value of 0.3. Shown in figure 2 are the experimental lifetimes for anthracene in supercritical CO_2 at 35 °C from a reduced density of 0.3 to 1.9. The results show a slight increase in τ_F with decreasing ρ_r. Using the data in the ρ_r region 0.8 to 1.9 only and plotting $\ln(k_F)$ vs $\ln(n)$ for anthracene, we obtain the plot shown in figure 5. As can be seen, the data is linear with a slope of 0.8. This value for the refractive index dependence is also somewhat lower than predicted. If we assume that the fluorescence quantum yield of anthracene will remain constant throughout the low density region, we can add the low density k_F values to the plot in figure 5. The low density k_F values thus calculated are in excellent agreement with the high density values, falling on the line with a slope of 0.8. It is quite apparent that the anomalous behavior observed for 9CA is not repeated with anthracene. Anthracene apparently has normal behavior with respect to fluorescence radiative rate theory.

Figure 3 shows the experimental lifetimes of 9,10-diphenylanthracene in supercritical CO_2 at 35 °C as a function of reduced density. The lifetime values show an odd dependence on reduced density. For reduced densities of 1.4 to 1.9, τ_F values increase slightly with increasing CO_2 density. For reduced densities of ~ 0.9 to 1.4, the τ_F values are fairly constant forming a plateau region. At reduced densities lower than 0.9, the τ_F values roll over and then decrease sharply with decreasing density. Fluorescence quantum yields for DPA in room-temperature solvents are reported between 0.9 and 1.0.[30] In a supersonic jet, the Φ_F value is 1.0.[30,31] If the k_F for DPA is to have approximate n^2 dependence, the Φ_F would have to vary in an odd manner that is significantly greater than the 10% variation suggested from the literature results. Assuming a quantum yield of 1.0, the plot of $\ln(1/\tau_F)$ vs $\ln(n)$ is shown in figure 6. In the high density region, the values of $1/\tau_F$ are unlike those observed for 9CA and anthracene. The values increase with decreasing density and refractive index, implying a negative refractive index dependence. In the low density region, the observed k_F results are similar to those of 9CA. $1/\tau_F$ tends to increase sharply with decreasing density. Obviously, the photophysical behavior of DPA in supercritical CO_2 is more complex than that observed for either 9CA or anthracene. It is quite apparent that accurate quantum yield measurements are required in order to understand DPA's photopysical behavior.

Bi-molecular quenching reactions in supercritical carbon dioxide. Simple bi-

Table 1.   Fluorescence Quantum Yields (Φ_F) and Lifetimes (τ_F) of the Probe Molecules in Non-polar Hydrocarbon Solvents and in a Supersonic Expansion.

Compound	Φ_F (solvent)	τ_F (ns)		Φ_F (vacuum)	τ_F (ns)
9CA[a]	1.0	14.5		1.0	28.0
DPA[b]	0.9	7.98		-	10.4
Anthracene[a]	0.31	5.3		0.67	21.5
Perylene[a]	0.94	6.4		0.97	9.3

[a] Taken form ref. 31; [b] Lampert, R. A.; Meech, S. R.; Metcalfe, J.; Phillips, D.; Schaap, A. A. P. *Chem. Phys. Lett.* **1983**, *94*, 137.

Table 2. Quenching Rate Constants k_q for the Probe-CBr$_4$ Systems Determined at Several Reduced Densities in Supercritical CO_2 at 35 °C.

Compound	Reduced Density				
	0.8	1.0	1.2	1.5	1.8
9CA	3.3×10^{10}	2.2×10^{10}	1.5×10^{10}	9.9×10^{9}	7.7×10^{9}
DPA	7.8×10^{10}	5.5×10^{10}	4.1×10^{10}	2.8×10^{10}	2.0×10^{10}
Anthracene	5.1×10^{10}	3.9×10^{10}	3.1×10^{10}	2.4×10^{10}	1.8×10^{10}
Perylene	6.2×10^{10}	5.2×10^{10}	4.4×10^{10}	3.3×10^{10}	2.4×10^{10}

units are in $M^{-1}s^{-1}$.

molecular reactions in supercritical fluids provide an efficient means for probing solute-solute interactions. Recently, the bi-molecular quenching reaction for the 9,10-bis(phenylethynyl)anthracene (BPEA) - carbon tetrabromide (CBr_4) system showed enhancements to the bi-molecular quenching rate constant of ~ 100% in the near-critical density region.[25] The enhancement was attributed to an increase in the local concentration of CBr_4 about the BPEA probe molecule. We have extended these studies with the quenching of four other related fluorophors with CBr_4 (Table I). The compounds differ in their photophysics with varying lifetimes and fluorescence quantum yields. All of the fluorescent molecules have been well studied in room-temperature solvents. The reactions with CBr_4 are known to be diffusion controlled and display linear Stern-Volmer kinetics.[34,35] Fluorescence lifetimes were measured for each of the probe molecules in the absence of and with varying concentrations of CBr_4 as a function of CO_2 reduced density. All of the quenching studies were performed in the ~0.8 to 1.9 reduced density region of supercritical CO_2 at 35 °C. Experimentally determined diffusion rate constants are determined with the application of the Stern-Volmer equation,

$$\tau_F^{\circ}/\tau_F = 1 + K_{SV} * [CBr_4] \tag{4}$$

$$K_{SV} = k_q * \tau_F^{\circ} \tag{5}$$

where τ_F° and τ_F are the lifetimes of the fluorescent probe without and in the presence of quencher, respectively, K_{SV} is the S-V quenching constant, and k_q is the quenching rate constant. Quenching rate constants are determined as a function of CO_2 reduced density and compared to diffusion rates predicted from the Debye equation,

$$k_{diff} = 8000RT/(0.3 * \eta) \tag{6}$$

where η is the solvent viscosity.

Shown in figures 7-10 are the plots of the experimental lifetimes for each of the fluorescent probes. In the absence of quencher, the density dependence of the lifetimes vary from probe to probe. 9CA, anthracene, and DPA were discussed in detail in the previous

5-13
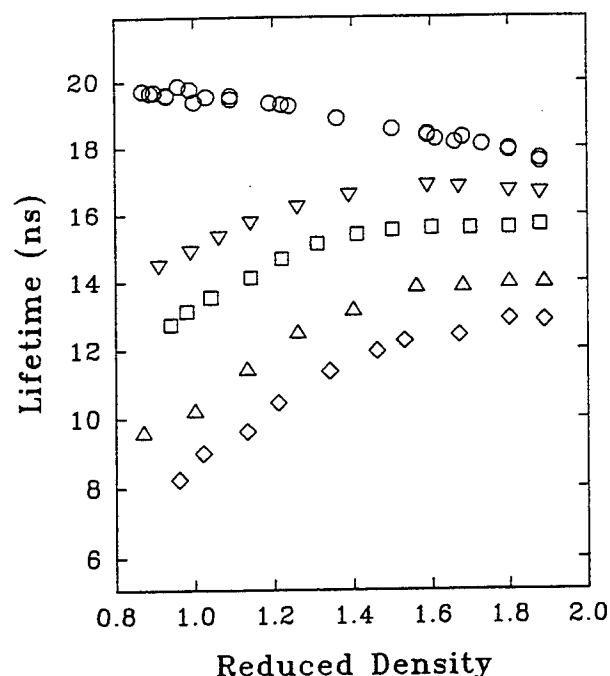
Figure 7.   Observed fluorescence lifetimes of 9-cyanoanthracene in supercritical CO_2 at 35 °C as a function of ρ_r at [CBr_4] of (top to bottom) 0, 0.51 x 10^{-3} M, 1.0 x 10^{-3} M, 2.0 x 10^{-3} M, and 2.9 x 10^{-3} M.
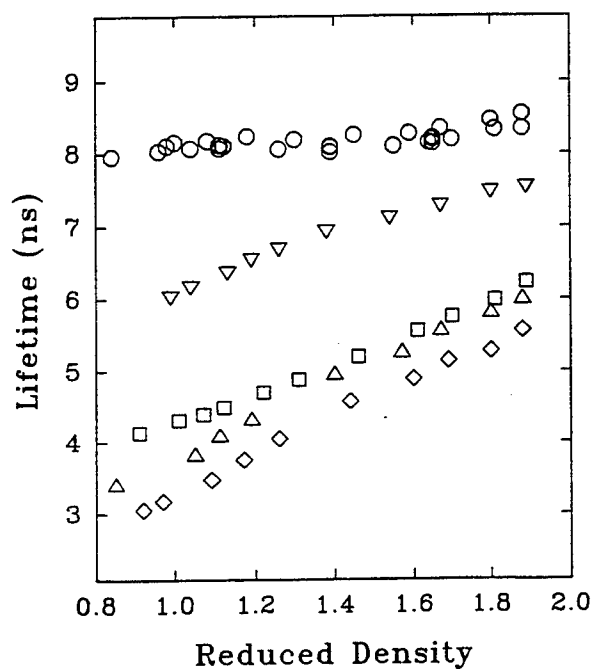
Figure 8.   Observed fluorescence lifetimes of 9,10-diphenylanthracene in supercritical CO_2 at 35 °C as a function of ρ_r at [CBr_4] of (top to bottom) 0, 0.73 x 10^{-3} M, 2.3 x 10^{-3} M, 2.6 x 10^{-3} M, and 3.3 x 10^{-3} M.
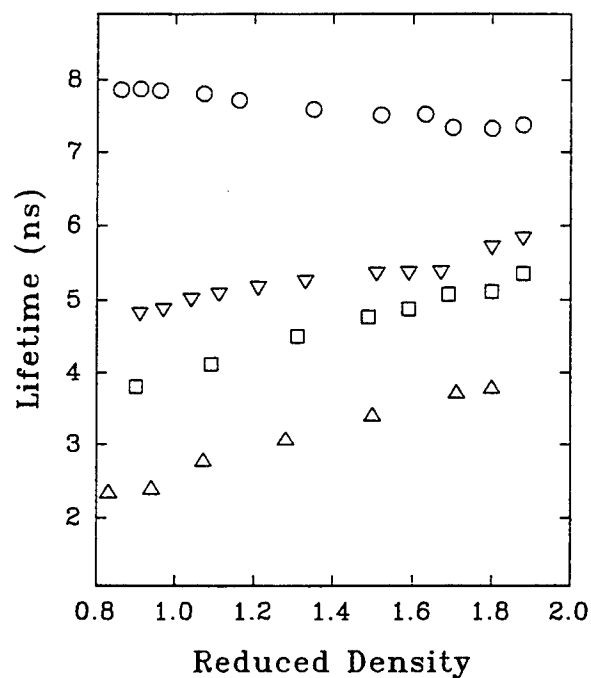
Figure 9.   Observed fluorescence lifetimes of anthracene in supercritical CO_2 at 35 °C as a function of ρ_r at [CBr_4] of (top to bottom) 0, 2.4 x 10^{-3} M, 3.0 x 10^{-3} M, and 6.2 x 10^{-3} M.
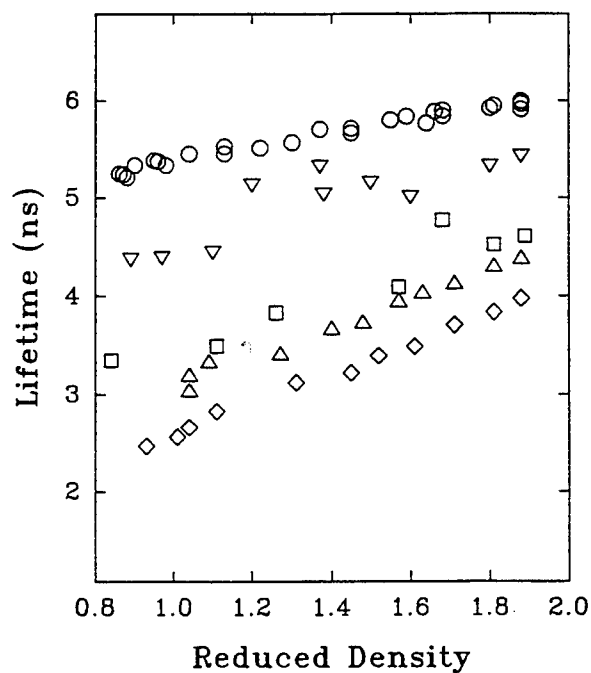
Figure 10.   Observed fluorescence lifetimes of perylene in supercritical CO_2 at 35 °C as a function of ρ_r at [CBr_4] of (top to bottom) 0, 0.94 x 10^{-3} M, 2.3 x 10^{-3} M, 3.2 x 10^{-3} M, and 4.3 x 10^{-3} M.

section. Perylene τ_F values decrease with decreasing reduced density in a linear fashion (figure 10). For each compound, addition of CBr_4 to the system results in quenching of the observed fluorescence lifetime. Higher CBr_4 concentrations result in greater quenching. At a constant CBr_4 concentration, the observed τ_F's will increase with increasing density. For anthracene, DPA, and perylene (figures 8-10), the increase in τ_F with increasing density is somewhat monotonic. For 9CA (figure 7), there is obvious curvature to the CBr_4 density dependent traces. The amount of increase in τ_F with increasing density is most significant in the near-critical density region. This is similar to the behavior observed for the $BPEA-CBr_4$ system.[25]

Quenching rate constants k_q are calculated from the experimental lifetimes as a function of CO_2 reduced density. The lifetime data was empirically fit with second or third order polynomial equations so as to allow interpolation and comparison of the data at any given density. Stern-Volmer plots as a function of reduced density were linear within experimental uncertainties. Table 2 reports k_q values of the four systems at selected reduced densities. At a reduced density of 1.9, k_q values are ~ 2.0 x 10^{10} $M^{-1}s^{-1}$ for anthracene, DPA, and perylene. 9CA is somewhat lower with a value of 7 x 10^9 $M^{-1}s^{-1}$. These quenching values are in reasonable agreement with liquid-like region results from other diffusion controlled systems in supercritical CO_2.[25] The k_q values tend to increase with decreasing density, however, each system appears to have its own dependence on density. A more dramatic comparison is shown in figure 11 where k_q for each system is normalized at $\rho_r = 1.9$ and plotted together. As can be seen, there are two apparent behaviors. 9CA and DPA both show more significant increases in k_q with decreasing ρ_r than anthracene and perylene. Interestingly, fluorescence quantum yields for 9CA and DPA are unity or near unity, while quantum yields for anthracene and perylene are 0.3 and ~ 0.94, respectively. When compared to the diffusion coefficients calculated from hydrodynamic theory, anthracene and perylene appear to be diffusion controlled throughout the CO_2 density region examined. 9CA and DPA, however, show an enhancement of ~ 25% in the near-critical region. The enhancement does not reach a maximum at a reduced density of 1.0, but continues to increase with decreasing reduced density all the way to $\rho_r = 0.8$. At $\rho_r = 0.8$, the increase in observed quenching rate constants is ~ 40 percent.

Clearly, the results for 9CA and DPA show an enhancement of the quenching rate constants above those predicted from theory. The enhancements, however, are significantly
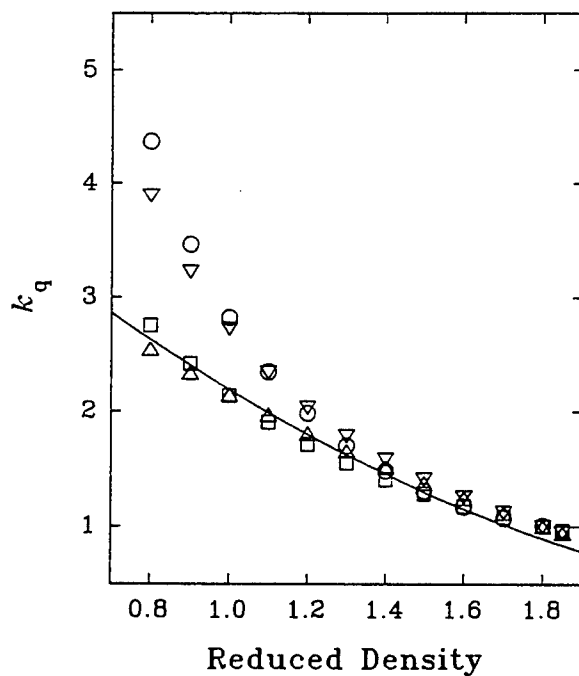
Figure 11.   Normalized quenching rate constants as a function of CO_2 reduced density at 35 °C (o, 9-cyanoanthracene; ▽, 9,10-diphenylanthracene; □, anthracene; and △, perylene).   The line represents the density dependence of the Debye diffusion rate constants.

lower than those observed for the BPEA-CBr$_4$ system. Considering that anthracene and perylene display normal diffusion controlled behavior when analyzed under the same conditions as 9CA and DPA, it is not surprising that these results are somewhat different. The results are still best explained as increases in the local concentration of the quencher with respect to the bulk concentration. Apparently, different reaction systems will have different responses. These results serve to reemphasize the complex nature of a supercritical fluid. Obviously, further work is required if models capable of accurately predicting complex chemical reaction behaviors are ever to be realized.

Conclusion

The studies of k_F and the bimolecular quenching reaction of the model fluorophors 9-cyanoanthracene, 9,10-diphenylanthracene, anthracene, and perylene as a function of supercritical carbon dioxide reduced density serve to demonstrate that supercritical fluids are complex solvent environments. Current understanding of these unique solvent systems is insufficient to accurately predict even the simplest of reaction systems. The fact that two of the three compounds used to investigate the k_F dependence on fluid density and refractive index showed anomalous results that can not easily be explained verifies this fact. The bimolecular quenching studies yielded similar results. Two of the four systems studied displayed unusual enhancements in the quenching efficiencies associated with the near-critical density region. The other two systems displayed normal diffusion controlled kinetics. No obvious explanation for why there exists this system dependence in observing the unusual properties of a supercritical fluid is available. Clearly, further investigation into this intriguing phenomenon is called for.

References

1. (a) Squires, T. G.; Paulaitis, M. E., Eds., *Supercritical Fluids, Chemical and Engineering Principles and Applications*; American Chemical Society: Washington, D.C., 1987.

2. Shaw, R. W.; Brill, T. B.; Clifford, A. A.; Eckert, C. A.; Franck, E. U. *Chem. Eng. News*, December 23, 1991, p. 26.

3. (a) Penninger, J. M. L.; Radosz, M.; McHugh, M. A.; Krukonis, V. J., Eds.,

Supercritical Fluid Technology; Elsevier: Amsterdam, 1985.

4. Savage, P. E.; Gopalan, S.; Mizan, T. I.;Martino, C. J.; Brock, E. E. *AIChE J.* **1995**, *41*, 1723.

5. Kim, S.; Johnston, K. P. *Ind. Eng. Chem. Res.* **1987**, *26*, 1206.

6. (a) Brennecke, J. F.; Tomasko, D. L.; Peshkin, J.; Eckert, C. A. *Ind. Eng. Chem. Res.* **1990**, *29*, 1682. (b) Eckert, C. A.; Knutson, B. L. *Fluid Phase Equil.* **1993**, *83*, 93.

7. Chialvo, A.; Cummings, P. T. *AIChE J.* **1994**, *40*, 1558.

8. (a) Kim, S.; Johnston, K. P. *AIChE J.* **1987**, *33*, 1603. (b) Bennett, G. R.; Johnston, K. P. *J. Phys. Chem.* **1994**, *98*, 441.

9. (a) Yonker, C. R.; Frye, S. L.; Kalkwarf, D. R.; Smith, R. D. *J. Phys. Chem.* **1986**, *90*, 3022. (b) Smith, R. D.; Frye, S. L.; Yonker, C. R.; Gale, R. W. *J. Phys. Chem.* **1987**, *91*, 3059. (c) Yonker, C. R.; Smith, R. D. *J. Phys. Chem.* **1988**, *92*, 235.

10. (a) Betts, T. A.; Bright, F. V. *Appl. Spectrosc.* **1990**, *44*, 1196, 1203. (b) Betts, T. A.; Zagrobelny, J.; Bright, F. V. *J. Supercritical Fluids.* **1992**, *5*, 48.

11. (a) Kajimoto, O.; Futakami, M.; Kobayashi, T.; Yamasaki, K. *J. Phys. Chem.* **1988**, *92*, 1347. (b) Morita, A.; Kajimoto, O. *J. Phys. Chem.* **1990**, *94*, 6420.

12. Sun, Y.-P.; Fox, M. A.; Johnston, K. P. *J. Am. Chem. Soc.* **1992**, *114*, 1187.

13. (a) Randolph, T. W.; Carlier, C. *J. Phys. Chem.* **1992**, *96*, 5146. (b) Carlier, C.; Randolph, T. W. *AIChE J.* **1993**, *39*, 876.

14. Howdle, S.; Bagratashvili, V. N. *Chem. Phys. Lett.* **1993**, *214*, 215.

15. (a) Sun, Y.-P.; Bunker, C. E.; Hamilton, N. B. *Chem. Phys. Lett.* **1993**, *210*, 111. (b) Sun, Y.-P.; Bunker, C. E. *Ber. Bunsen-Ges. Phys. Chem.* **1995**, *99*, 976.

16. (a) Wu, R.-S.; Lee, L. L.; Cochran, H. D. *J. Supercrit. Fluids* **1992**, *5*, 192. (b) Cochran, H. D.; Johnson, E.; Lee, L. L. *J. Supercrit. Fluids* **1990**, *3*, 157. (c) Jonah, D. A.; Cochran, H. D. *Fluid Phase Equilib.* **1994**, *92*, 107.

17. (a) Petsche, I. B.; Debenedetti, P. G. *J. Chem. Phys.* **1989**, *91*, 7075. (b) Tom, J. W.; Debenedetti, P. G. *Ind. Eng. Chem. Res.* **1993**, *32*, 2118.

18. (a) Debenedetti, P. G.; Mohamed, R. S. *J. Chem. Phys.* **1989**, *90*, 4528. (b) Chialvo, A. A.; Debenedetti, P. G. *Ind. Eng. Chem. Res.* **1992**, *31*, 1391.

19. Chialvo, A. A. *J. Phys. Chem.* **1993**, *97*, 2740.

20. O'Brien, J. A.; Randolph, T. W.; Carlier, C.; Shankar, G. *AIChE J.* **1993**, *39*, 1061.

21. (a) Kurnik, R. T.; Reid, R. C. *Fluid Phase Equil.* **1982**, *8*, 93. (b) Kwiatkowski, J.; Lisicki, Z.; Majewski, W. *Ber. Bunsen-Ges. Phys. Chem.* **1984**, *88*, 865.

22. Brennecke, J. F.; Tomasko, D. L.; Eckert, C. A. *J. Phys. Chem.* **1990**, *94*, 7692.

23. Combes, J. R.; Johnston, K. P.; O'Shea, K. E.; Fox, M. A. in *Supercritical Fluid Science and Technology*; Johnston, K. P.; Penninger, J. M. L., Eds., American Chemical Society: Washington, D.C., 1989, p. 31.

24. (a) Kazarian, S. G.; Gupta, R. B.; Clarke, M. J.; Johnston, K. P.; Poliakoff, M. *J. Am. Chem. Soc.* **1993**, *115*, 11099. (b) Gupta, R. B.; Combes, J. R.; Johnston, K. P. *J. Phys. Chem.* **1993**, *97*, 707.

25. Bunker, C. E.; Sun, Y.-P. *J. Am. Chem. Soc.* **1995**, *117*, 10865.

26. (a) Zagrobelny, J.; Betts, T. A.; Bright, F. V. *J. Am. Chem. Soc.* **1992**, *114*, 5249. (b) Rice, J. K.; Niemeyer, E. D.; Dunbar, R. A.; Bright, F. V. *J. Am. Chem. Soc.* **1995**, *117*, 5832.

27. (a) Roberts, C. B.; Zhang, J.; Brennecke, J. F.; Chateauneuf, J. E. *J. Phys. Chem.* **1993**, *97*, 5618. (b) Roberts, C. B.; Chateauneuf, J. E.; Brennecke, J. F. *J. Am. Chem. Soc.* **1992**, *114*, 8455.

28. Roberts, C. B.; Zhang, J.; Chateauneuf, J. E.; Brennecke, J. F. *J. Am. Chem. Soc.* **1993**, *115*, 9576.

29. (a) Zagrobelny, J.; Bright, F. V. *J. Am. Chem. Soc.* **1992**, *114*, 7821. (b) Zagrobelny, J.; Bright, F. V. *J. Am. Chem. Soc.* **1993**, *115*, 701.

30. Lampert, R. A.; Meech, S. R.; Metcalfe, J.; Phillips, D.; Schaap, A. P. *Chem. Phys. Lett.* **1983**, *94*, 137.

31. Hirayama, S.; Iuchi, Y.; Tanaka, F.; Shobatake, K. *Chem. Phys.* **1990**, *144*, 401.

32. Sun, Y.-P.; Fox, M. A. *J. Phys. Chem.* **1993**, *97*, 28239.

33. Rice, J. K.; Niemeyer, E. D.; Bright, F. V. *J. Phys. Chem.* **1996**, *100*, 8499.

34. (a) Birks, J. B. *Photophysics of Aromatic Molecules*; Wiley-Interscience: London, 1970. (b) Birks, J. B. In *Organic Molecular Photophysics, Vol. 2*; Birks, J. B., Ed.; Wiley: London, 1975, Chapter 9.

35. Leite, M. S. S. C.; Naqvi, K. R. *Chem. Phys. Lett.* **1969**, *4*, 35.

SIMULATION OF ANTI-JAMMING GPS ARRAYS USING FINITE ELEMENT SOFTWARE

Mark Casciato
Graduate Research Assistant
Radiation Laboratory
Dept. of Electrical Engineering and Computer Science

University of Michigan
1301 Beal Ave.
Ann Arbor, MI 48109-2122

SIMULATION OF ANTI-JAMMING GPS ARRAYS USING FINITE ELEMENT SOFTWARE

John L. Volakis[1] and Mark Casciato[2]
[1]Professor   [2]Graduate Research Assistant
Radiation Laboratory
Dept. of Electrical Engineering and Computer Science
University of Michigan

Abstract

Two sophisticated anti-jam antennas explicitly designed for the Global Positioning System (GPS) satellite communications were simulated and analyzed using the University of Michigan finite element software FEMA-PRISM. The software was upgraded to allow simulation of these unique GPS antenna arrays which consisted of a reference element and an array of six anti-jam auxiliary elements that produce pattern nulls at the direction of the incoming jamming signal. The goal of the study was to examine the effectiveness of current state-of-the-art software for simulating complex antenna arrays and to generate accurate volumetric patterns (incorporating mutual coupling, losses, dispersion, etc.) of the actual array to be used in a hardware simulator when evaluating the anti-jam performance of the antenna array under test. Our study demonstrated that the University of Michigan finite element code FEMA-PRISM with (minor modifications) was capable of analyzing the two GPS antennas without compromising their geometrical features in performing the numerical analysis. For one of the arrays, the antenna element was a slot backed by a triangular cavity whereas for the second array the basic element was a dual patch to accommodate radiation at the L1 (1575 MHz) and L2 (1227 MHz) GPS bands. Patterns and input impedance curves were generated for each antenna and these were validated with reference data whenever possible. With the availability of the simulation models and discrete geometry data for each antenna, future efforts will focus on evaluating the performance of the arrays under various jamming conditions and realistic GPS operational scenarios.

SIMULATION OF ANTI-JAMMING GPS ARRAYS USING FINITE ELEMENT SOFTWARE

John L. Volakis and Mark Casciato

Introduction

Over the past 7 years, the University of Michigan has been developing computational methods and related computer codes for the analysis of antennas on complex platforms. Antennas are the "eyes" and "ears" of all communication devices, including radar, microwave imaging systems, intelligent vehicle guidance systems and for wireless communication devices for cellular, personal communication systems(PCS), global positioning system (GPS) and lacal area networks (LAN) applications. The latter has been a major driver toward the design, development and deployment of suitable antennas which are efficient and satisfy the requirements of personal, military and commercial systems.

In the past, antenna design was based on a few well known geometrical shapes such as patches, horns, reflectors, spirals, log-periodic, Yagi-Uda(used for TV reception), wire dipoles (typically used for car radios) and wire loops. Some of these are known to be broadband (for example, log-periodic) and others are known to have very narrow bandwidth (for example, the patch). Designers usually worked with these classic shapes and chose their geometrical parameters to satisfy the frequency, bandwidth and gain pattern requirements. Deviation from the classical shapes requires substantial investment in experimental and analytical research and development. Increasing and specialized requirements on pattern, bandwidth and gain require new antenna designs as well as feed networks. Indeed, the cost of developing new antenna designs and analysis by experimental means is prohibitive and over the past few years there has been an impetus to develop antenna analysis and design software. At the University of Michigan we developed a unique code referred to a FEMA-PRISM for conformal antenna analysis[1]. This code is the first to be suitable for simulating antennas on doubly curved platforms. Moreover, it is based on a new hybrid finite element methodology which permits the analysis of antennas of varying shape and material composition provided the geometrical data are supplied in a certain form. Additional information about FEMA-PRISM and its formulation can be found the references [2-11]. A brief description of the University of Michigan finite element analysis codes is given in the Appendix.

Scope of the Work

The scope of our work over the past two months has been the application and further development of FEMA-PRISM to two GPS antennas with anti-jamming auxiliary antenna elements in support of the Wright Laboratory anti-jamming antenna array simulator[12]. This hardware simulator evaluates the anti-jamming performance of various antenna arrays (over a frequency band and as a function of reception angle). A typical antenna element along with anti-jamming array elements is depicted in Figure 1. Presently, the simulator assumes that each of the antenna elements is radiating as isotropic source (i.e., it is assumed to radiate equally in all directions), whereas in practice
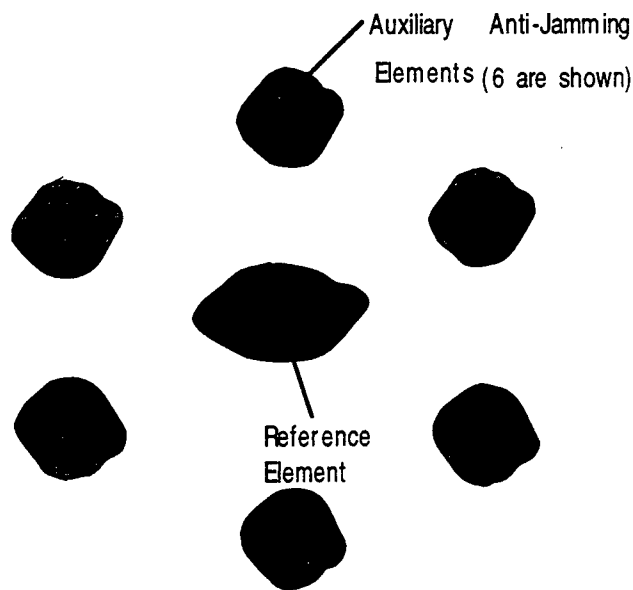
Figure 1. Illustration of the anti-jamming array. The reference element is the main receiver and the excitation of the auxiliary elements are adjusted to generate a null at the direction of the jamming signal.
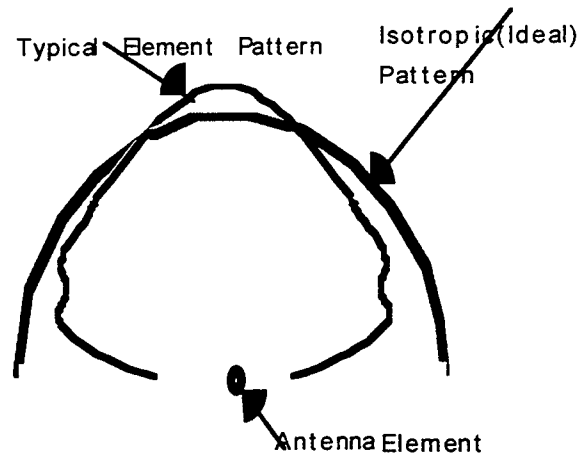
Figure 2. Comparison of radiation patterns due to an isotropic source and that of an typical actual element.

the antenna pattern can deviate substantially from the isotropic pattern conditions as illustrated in Figure 2. The actual pattern depends on the shape and geometry of the element as well as the shape of the platform (aircraft, missile and so on). Furthermore, the anti-jamming element patterns and their coupling with the reference element (or among themselves) present additional variables which must be considered in the performance of the array in the presence of jamming signals. Typically, the anti-jamming elements are used to generate a canceling signal (nulling) at specific directions (the jamming signal directions) using adaptive signal processing algorithms. This cancellation must be accomplished with extreme accuracy and depth using information related to the phase and amplitude of the received signal at the various array elements. The pattern of the individual elements and their coupling play an important role in rejecting the jamming signals while preserving the array's capability to receive information from other directions. It turns out that accurate information relating to the dielectric constants of the material surrounding the antenna array elements is also needed. Indeed, even a small variation of the dielectric constants as a function of frequency can compromise the performance of the nulling algorithm. This is due to different surface wave signal propagation speeds caused by the changing dielectric constant. As a result, the assumed linear phase variation with respect to frequency no longer holds. The phenomenon is referred to as signal dispersion and compromises the performance of the adaptive nulling algorithm. To accurately model antenna performance characteristics, it is necessary to account for the specific element pattern shape as well as coupling and dispersion effects caused by the dielectric used in place of the simulator's ideal antenna model (i.e. uncoupled, linear and isotropic elements). To do so, it is necessary to model the entire antenna structure including its detailed geometrical and material properties and to include coupling and platform curvature effects. The FEMA-PRISM code allows for the precise modeling of all

these antenna attributes and can simulate the antenna array patterns under any excitation. During the past two months the following tasks were carried out:

1. Modeled the two GPS antenna configurations using FEMA-PRISM and examined their performance around the L1 and L2 GPS bands (1227.6 MHz and/or 1575.42 MHz). The actual geometry generation, code modifications and runs were carried out by Mark Casciato who is now pursuing his Ph.D. degree at the University of Michigan. Prof. Volakis worked with Mr. Casciato in defining the analytical model, obtaining the geometry data from the USAF suppliers, data analysis, code upgrades and in planning the analysis work.

2. Wrote a 30 pp. report which provides an overview of the finite element method for antenna simulations and how the method is implemented into FEMA-PRISM. This document was also provided with five recent journal papers describing with greater detail the working principles of FEMA-PRISM and previous finite element codes which lead to the development of FEMA-PRISM. (Completed by Prof. Volakis)

3. Wrote a 30 pp. document on the basic principles of the finite element method at a level to be read by engineers unfamiliar with numerical methods and who are interested in an intelligent use of the FEMA-PRISM code. (Completed by Prof. Volakis)

4. Updated the FEMA-PRISM manual to incorporate lumped load modeling and interface with SDRC-Ideas used for meshing the antenna surface. (Done by Mr. Casciato)

5. The source code along with interfaces to MATLAB for plotting, meshing packages as well as pre- and post-processing were left with the Wright Laboratory. As is the case with all software technologies, FEMA-PRISM will need to be updated continuously to accommodate new antennas and to increase its usability to a wider community base.

Below we describe the analysis carried out for the two unique GPS antennas.

Lockheed CRPA2 Anti-Jamming GPS Array

The layout of this anti-jamming array is similar to that in Figure 1. Six anti-jamming array elements are placed around the reference element. The geometry of the anti-jam element is shown in Figure 3 based on the dimensions provided by Mr. Pellet of Lockheed. It consists of a triangular cavity which is divided in two chambers using a metallic separator with an iris placed in the middle of the separator to allow excitation of both cavity chambers. Two capacitors are used for controlling the cavity Q and thus permitting resonance at L1 and L2 bands. The location and layout of the capacitors were of sensitive nature and had to be provided through the F117 U.S. Air Force System Program Office (SPO). Unfortunately, this information was not provided till one week prior to the end of the Summer Faculty Appointment. Thus we initially proceeded to model the Lockheed CRPA2 without the capacitor elements with the intention of adding these at a later stage. Lack of the capacitor elements will cause the cavity to resonate at different frequencies and thus much of the geometry meshing and simulation can be done without a knowledge of the capacitor elements.
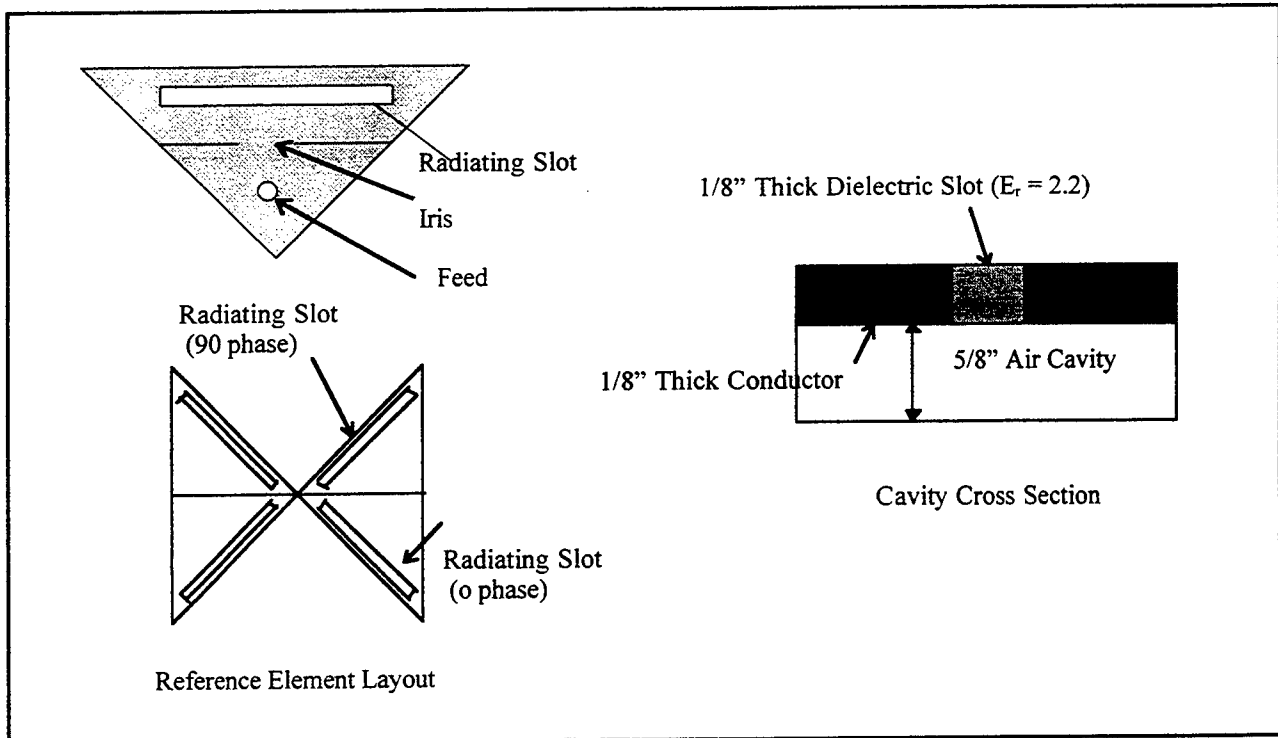
Figure 3. The Lockheed CRPA2 triangular auxiliary antenna element. Four of these constitute the reference element arranged as shown.

The cavity was fed by a coaxial cable which was modeled by a probe feed at the shown location and radiation occurred through the straight slot located near the edge of the cavity opposite to the right angle. When the cavity is at resonance, strong fields are generated within the cavity and under the slot implying maximum radiation at the corresponding frequencies. When the cavity is not at resonance, the modal cavity fields are highly attenuated before they reach the slot location. As a result, the electric field across the slot is small when the cavity is not at resonance and little or no radiation occurs at the off-resonance frequencies. The reference element consisted of four cavities arranged so that the slots form an X pattern. Since the slots are fed in quadrature with respect to the adjacent element/slot, the resulting radiation is circularly polarized whereas the excitation of the auxiliary anti-jam elements is controlled by the adaptive nulling network to generate a null in the direction of the incoming jamming signal. Thus, as a start we decided to provide a complete characterization of the single triangular cavity since the entire array, including the reference element can be modeled by repeating this element with different excitations. We will refer to the element shown in Figure 3 as the basic element of the Lockheed CRPA2.

In modeling the anti-jamming GPS antennas, the following steps were taken:

1. Translate actual model to one suitable for FEMA-PRISM simulation. This is usually an easy step and amount to certain minor simplifications. However, some difficulty was encountered due to data sensitivity issues which took some time to identify and resolve.

2. Generate a surface mesh using the FEMA-PRISM's internal meshing package or an external mesher. The latter is necessary for complex antenna configurations and for our case, we used a commercial package by SDRC-

Ideas. A translator was written for this package to reformat the data-sets. Given the surface mesh of the antenna element, FEMA-PRISM generates the volume mesh by growing prismatic elements away from the surface.

3. Post-process mesh to enforce boundary conditions and to introduce lumped elements such as resistors, capacitors, shorting pins, via holes and feed geometry. These are of course essential for the correct operation of the antenna. For the case of the Lockheed CRPA2, a special preprocessor was written to handle the conductor thickness. As shown in figure 3, the slot was formed by a 1/8" thick metal whose thickness is essential in controlling the resonance frequency. To ensure the correct operation of the PRISM code, a single slot was first modeled which was 2.75" long, 1/8" wide and was placed in a 5/8" deep rectangular cavity as in the Lockheed CRPA2 configuration. A dielectric having a relative permittivity of $\varepsilon_r =2.2$ was also placed in the slot. The results of PRISM were compared to a reference code and agreed to within 2.5% percent as shown in Figure 4. For this simulation we used 23876 non-zero edges (out of 33410) whereas the reference FEMA_BRICK code used 11948 edges. About 3500 iterations were required by FEMA_PRISM to achieve convergence.

4. Execute FEMA_PRISM code for the specified antenna geometry and excitation to obtain input impedance, near zone fields, radiation patterns, coupling effects and other parameters of interest.

The surface mesh corresponding to the antenna element in Figure 3 was generated using SDRC-Ideas and is illustrated in Figure 5. Typical results at a frequency of 1575 Mhz are displayed in Figure 6 which include two near zone plots for each of the electric field components that couple to the slot and the far zone pattern of the element. Note that the horizontal electric field component is the most dominant in the slot and is responsible for radiation. Note also that these plots were generated without the cavity separator. The convergence of the FEMA-PRISM biconjugate gradient iterative solver was extremely good for this application and required just 278 iterations for a system of about 30,000 degrees of freedom. This is close to the theoretical convergence rate of about 1/100 of the number of unknowns implying a well conditioned system. Such system characteristics are indicative of high accuracy. Also, as demonstrated by the near zone fields, the slot is excited by the cavity field.

Having successfully completed the element simulation with FEMA-PRISM we could now proceed with the array analysis, including the jamming elements with the goal of obtaining accurate patterns which include coupling effects and inherent features due to the antenna element geometry. However, given that critical information were still not available from Lockheed, we instead proceeded to model another GPS antenna candidate which employs a set of two stacked patches as described next.
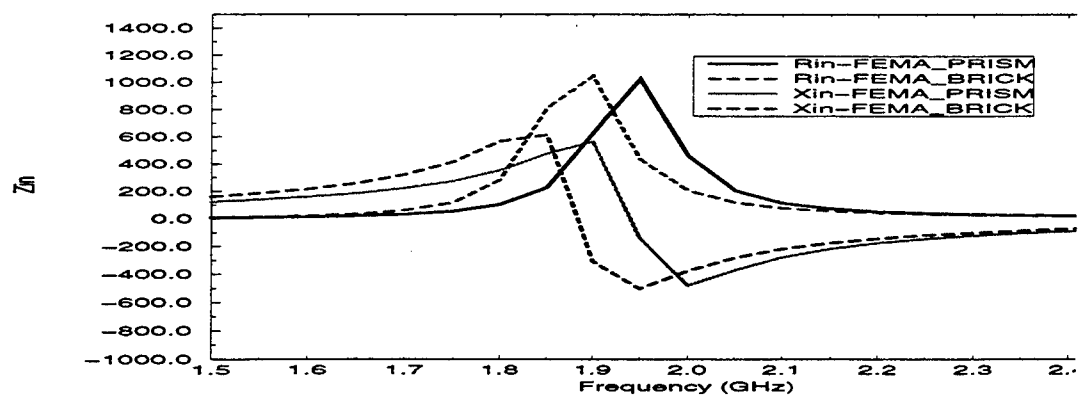
Figure 4. Slot resonance results. Comparison of FEMA-PRISM and FEMA-BRICK
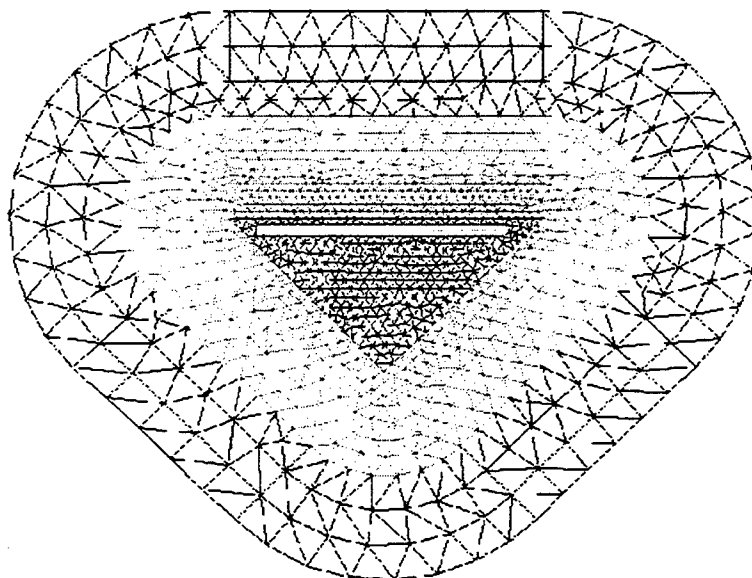
Figure 5. Surface Mesh of the Lockheed CRPA2 basic element

TRIANGULAR CAVITY SLOT

NEAR FIELDS

NORMALIZED RADIATION PATTERN
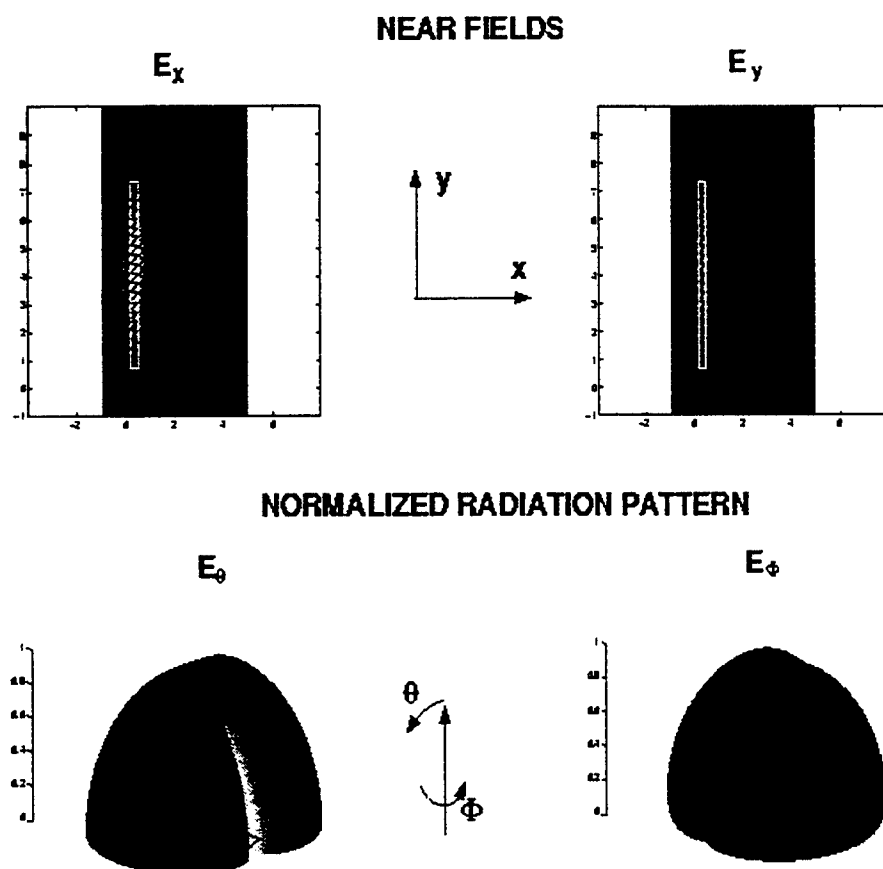
Figure 6. Results of FEMA-PRISM for the Lockheed CRPA2. Near zone fields and radiation pattern.

Ball Communications Dual Patch Anti-Jamming GPS Array

The Ball Communications dual patch array is a similar configuration to Figure 1. The Ball elements however consist of a stacked patch configuration as shown in Figure 7. In this configuration the smaller top patch is for the L1 (1574.42 Mhz) frequency with the inner-cavity patch for L2 (1227.6 MHz). A coaxial cable, routed through a hole in the L2 patch, drives the L1 patch with the L2 patch being parasitic. Again the coax is modeled as a current probe feed.
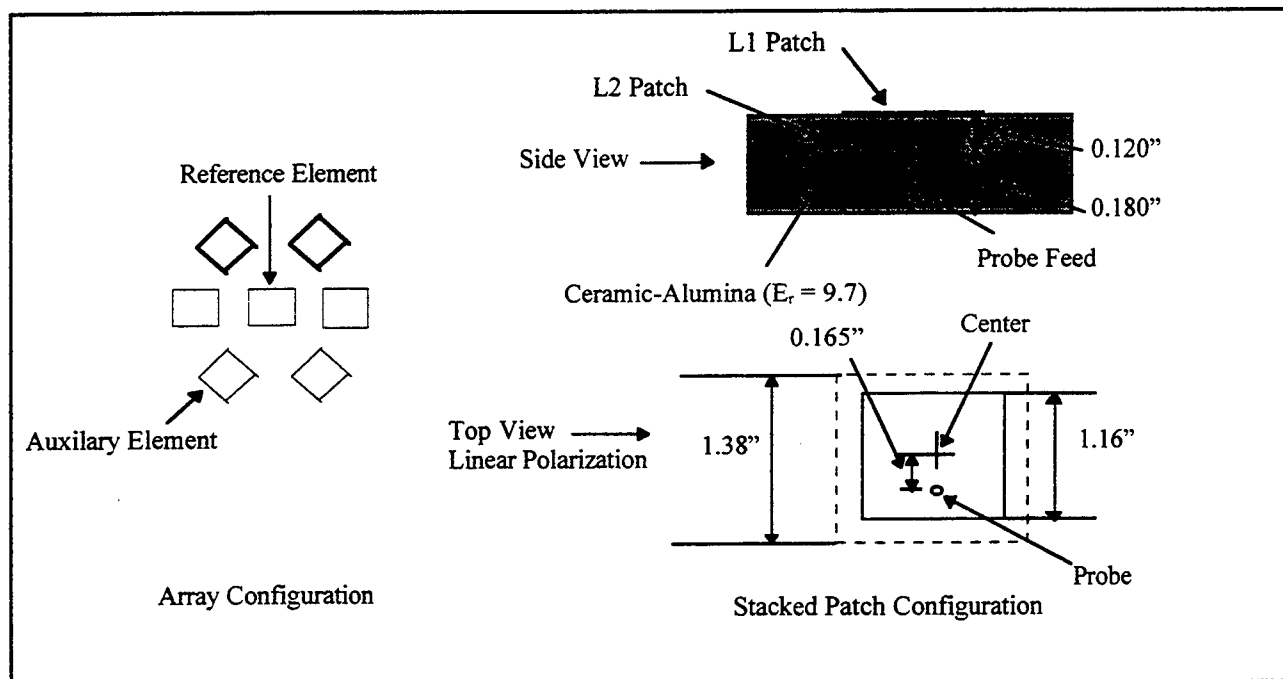
Figure 7. Ball Dual Patch Antenna. Array and Stacked Patch Configuration.

The array layout is also shown in Figure 7. As was the case with the Lockheed antenna detailed information on the Ball antenna was not available at this time. Due to this, a stacked patch configuration as shown in Figure 7 was used[13]. Note that the probe's offset location from center guarantees linear polarization. A second probe, with a 90^0 phase shift from the one shown, placed right/left of patch center will create circular polarization. The following task were carried out in connection with the Ball array.

1. To account for the larger L2 patch and hole, code modifications were made to permit a varied patch and hole size in specified layers. To validate these changes a single reference patch was generated and again compared to FEMA-BRICK data[13]. The results showed the antenna element resonances at the proper L1 and L2 locations as illustrated in Figure 8. However, in comparison to the FEMA-BRICK data[3] a difference in absolute impedance levels was observed. A possible explanation for this difference is that while the hole was modeled as perfectly square in FEMA-BRICK, for the FEMA-PRISM model the central patch along with the patch itself had trimmed corners. Figures 9a,c show the normal directed fields directly under each patch overlaid with the patch outline. As can be seen in Figure 9c the fields under the L2 patch show perturbations due to the trimmed

corners which will have an effect on the impedance levels. Also absolute impedance levels are difficult to predict. Figures 9b,d show the respective far-field patterns of the L1 and L2 frequencies. For the results in Figure 9, the linear system was approximately 71,000 unknowns and convergence was achieved in 1800-4000 iterations depending on the frequency of calculation. Typically, the convergence is slower at the resonance frequency because of the higher field singularities. The runs were executed on an SGI Indigo 2 and the CPU time per iterations was 2 seconds.

2. Generate a surface mesh using SDRC-Ideas of the full array, time permitting

3. Execute FEMA_PRISM for the full array geometry including mutual coupling factors, time permitting. Due to the limited information available the stacked patch configuration from [13] will be used to model the Ball array. This will be a proof of concept showing FEMA_PRISM's ability to model all array interactions including nulling produced by the auxiliary elements.
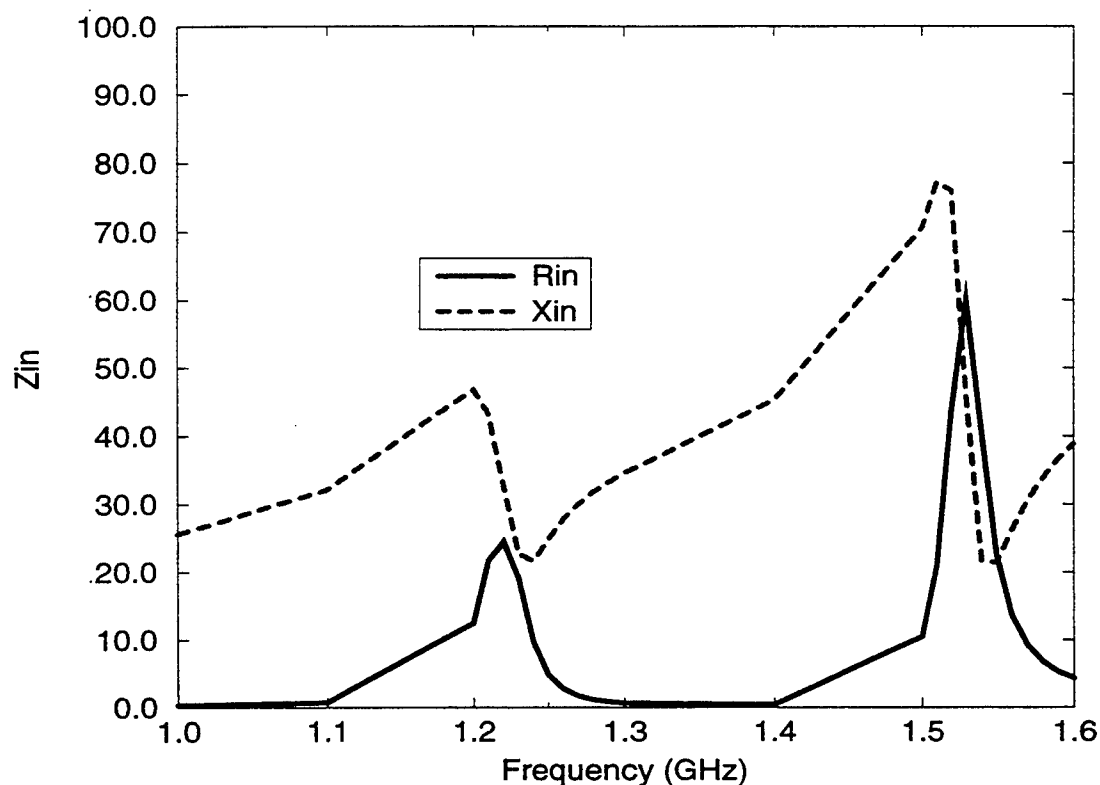
Figure 8. Input impedance of the dual patch Ball Communications element as a function of frequency. Note that the spikes in the plot indicate the frequencies at which the patch is resonant.
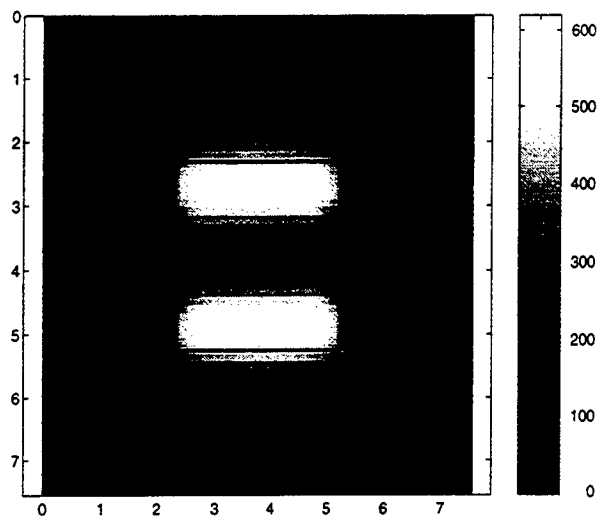
Figure 9a. Substrate Fields at L1 band.
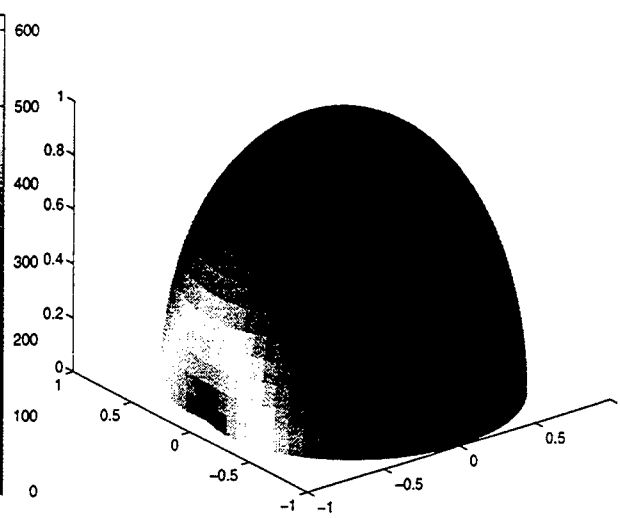
Figure 9b. Far-Field Pattern at L1 band.
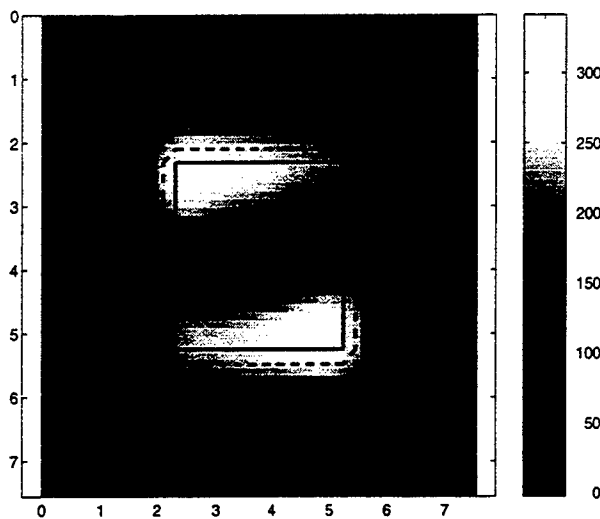
Figure 9c. Substrate Fields at the L2 band.
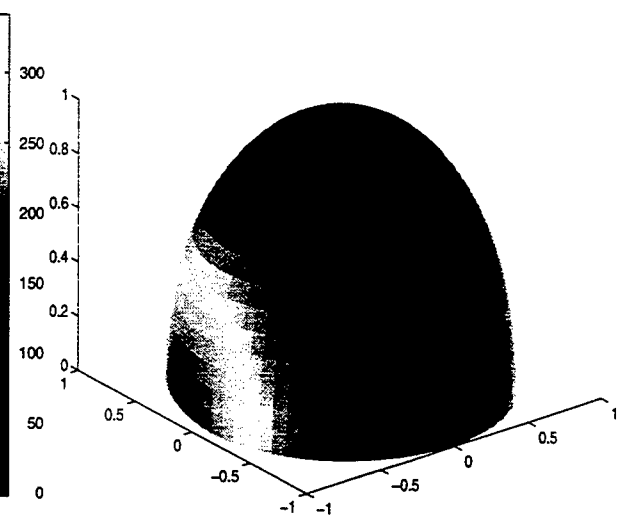
Figure 9d. Far-Field Pattern at the L2 band

FEMA PRISM Code Improvements and Upgrades.

Several changes were made to FEMA_PRISM and supporting codes to allow modeling of the anti-jam arrays. In addition many of these changes are applicable to other problems and enhance the abilities of FEMA_PRISM to handle future antenna configurations. All code modifications were made with the user in mind. With few exceptions, most changes are transparent to the user and are simply a question of modifying the FEMA_PRISM input file, MainInput. Also, where necessary, the changes were validated with simple examples and compared to reference data.

1) PreProcessing:

The ability to process CAD files is essential for a code to analyze various antenna platforms. CAD packages such as SDRC-Ideas, Hypermesh, etc., allow the user to easily generate various surface meshes of high quality. Output files contain information such as node location, connectivity, material identifiers, conducting pin location, etc. With this in mind a preprocessor was written to operate on an SDRC-Ideas universal file and convert it to a FEMA_PRISM compatible mesh. Note that other CAD files use similar methodology to describe mesh information. As other CAD packages are evaluated it is simply a matter of modifying the preprocessor to read in the new file format. Everything else is taken care of automatically. A sample preprocessor input file with comments is shown below.

```
            !Input Universal File Name.
sltele1.unv
            !Output FEMA_PRISM File Name.
test1.mesh
1           !Node Convention (1)LHN,(2)RHN.
-45         !Object Rotation (Degrees).
2           !Cad File Units.
            !Meters(1), cm(2), mm(3), Feet(4), inches(5)
0           !Iris for Slot Dual Cavity: no(0), yes(1)
```

As can be seen the preprocessor reads in an Ideas universal file, and outputs a FEMA_PRISM compatible mesh. The user specifies the universal input file node convention (FEMA_PRISM internally uses RHN convention); degrees to rotate the mesh if desired; and the universal file units (FEMA_PRISM uses cm). The last option is experimental to read in iris information for the Lockheed CRPA2, and should be left at 0 for most applications.

For a final visual check, the preprocessor writes out the FEMA_PRISM mesh file and also generates it's own universal file, chkslt.unv, with each material parameter given a separate color so that the user can visually check the final mesh and mesh parameters in the CAD package. Note that in the universal file nodes are written first, followed by the connectivity tables. Next, the conducting (grouped in Ideas by the user) and the iris pin nodes are listed, if necessary. The following material convention was used:

Material #	Material Type
1	air
2	absorber
3	dielectric under conductor(surface patch)
4	dielectric (exposed portion of dielectric)

2) Internal Rectangular mesh generator:

FEMA_PRISM contains an internal mesh generator which generates a simple rectangular mesh. The option was added in the MainInput file to generate a rectangular slot instead of a patch.

3) Conducting Pins:

An important addition to FEMA_PRISM is the ability to define conducting edges or pins in the cavity volume. In the case of FEMA_PRISM, these are vertically directed pins i.e., pins normal to the mesh surface. The surface position of these pins are identified in the Ideas universal file and automatically added to the FEMA_PRISM input surface mesh file. Along with inner-cavity conducting patches, these pins allow the user to define a finite thickness of the conducting patch. For resonant cavities, a finite conductor thickness as opposed to a zero thickness conductor, makes for a more accurate and stable numerical solution. These pins are added to layers in the cavity and must therefore conform to the user defined layers in the cavity. To add conducting pins in any layer, the user in the FEMA_PRISM input file, specifies the number of layers that will contain conducting pins, and then lists the pin layers.

4) Conducting Layers:

The ability to specify conducting layers or inner-cavity patches is another important addition to FEMA_PRISM. Coupled with conducting pins, these allow the user to define a conductor thickness. In addition, conducting layers allow for modeling stacked patch antenna configurations. As with conducting pins, the user specifies in the input file the number of conducting layers and lists the appropriate layers. At the top of each layer, a duplicate of the surface conducting patch is be inserted. Again the specified layers must conform to the user defined cavity layers.

5) Expanded Layers/Holes:

With the ability to expand conducting patches in any layer and to insert a hole in these patches FEMA_PRISM can now analyze stacked patch antenna configurations. The expanded layer must be one of the conducting layers defined previously. In the FEMA_PRISM input file, the user simply defines the number of segments to expand the patch. For each segment, the edges radiating from the conducting patch along with the edges connecting these radiating edges are defined as conducting. For patch holes, the layer containing the hole must be one of the conducting layers defined previously. The user specifies the layer containing the hole, the number of holes in that layer, the hole locations (i.e., node where the holes will be centered is usually the same location as the feed probe), and the number of segments/elements to expand the square hole. All holes in a given layer will be expanded by the same amount. For the first expanded segment, edges radiating from the specified nodes will be defined as non-conducting. For each additional segment the hole is expanded, non-conducting hole edges defined in the previous iteration will be

connected by non-conducting edges, and then edges radiating from these new non-conducting edges will be defined as non-conducting.

6) Iris:
For the Lockheed CRPA2 triangular cavity the option to add a conducting wall to the cavity was added. The surface nodes where the wall is placed is identified again in the Ideas universal file and transferred to the FEMA_PRISM mesh file. The wall is then automatically grown between the specified nodes. If a vertical iris is desired in the wall the opening must be at least 2 segments wide or the wall will extend the breadth of the cavity with no opening.

7) QMR option:
In addition to the currently available BiCG iterative solver, the option of a QMR solver was added. This solver converges quasi-monotonically as opposed to the oscillating convergence of BiCG. While the QMR solver was fully tested for accuracy, due to time limitations, it wasn't optimized for speed and further work is needed. The user specifies in the input file whether BiCG or QMR is preferred.

8) Diagonal Preconditioner:
A diagonal preconditioner was added to FEMA_PRISM to increase the rate of convergence. It was found that while a significant improvement was seen when QMR was specified, BiCG reacted unfavorably to this type of pre-conditioning. Consequently, preconditioning is only implemented internally when QMR was specified. Note that the pre-conditioned QMR converges at approximately the same rate as the unconditioned BiCG.

Conclusions and Future Directions

The most critical aspect of this Tour has been the identification of a situation, where a new software technology can have a major impact on our ability to develop real world anti-jamming array simulations without resorting to extensive measured databases that are likely to be impractical and unsuitable for array optimization and improvements. Thus, it is important to note that the antenna characterization tasks were carried out by FEMA-PRISM methodically and without difficulty. This clearly demonstrates the code's capability to handle these GPS communication antennas as well as other configurations used on GPS-guided munitions. The latter are mouted on doubly curved platforms and a unique feature of FEMA-PRISM is its capability to readily handle arrays on nonplanar surfaces and to account for coupling and element interactions in the presence of complex materials. Presnetly, the array simulation models used in conjunction with the Wavefront Simualtor (AWFS) facility are based on "ideal" isotropic sources, and thus the FEMA-PRISM code can become an invaluable and possibly necessary addition to the AWFS. It can provide pattern data for the actual array configuration and this data can be directly fed to the AWFS for real-time evaluation of the array performance under various jamming conditions. That is, the array analysis software can be installed on a dedicated computational workstation and integrated with the AWFS for a performance evaluation under conditions which are nearly identical to the real-world situation. The risk for the success of this hardware-in-the-loop simulator is indeed very low and the cost is limited to the specialization of the software for such an application. Before, proceeding with the implementation of such a computational workstation, it is necessary

to make further utilization of the code for various arrays

excersize the code in connection with a variety anti-jamming scenarios

make extensive comparisons with measurements

simulate protype situations of the integrated workstation platform

develop suitable I/O interface to the code for ease of use as part of the AWFS

These tasks are, of course, quite substantial since they outline a project that will be of major benefit to the Air Force in terms of simulation accuracy and cost-reduction. The Summer Faculty Research Continuation Grant appears to be a most prudent way to continue toward the completion of these tasks. Of course, the realization of these tasks and goals cannot be achieved without the commitment of substantially more resources.

Appendix: UNIVERSITY OF MICHIGAN FINITE ELEMENT CODES

Over the past 10 years we have witnessed an increasing reliance on computational methods for the characterization of electromagnetic problems. Although traditional integral equation methods continue to be used for many applications, one can safely state that in recent years the greatest progress in computational electromagnetics has been in the development and application of partial differential equation(PDE) methods such as the finite difference-time domain and finite element (FEM) methods[3], including hybridizations of these with integral equation and high frequency techniques. The major reasons for the increasing reliance on PDE methods stem from their inherent geometrical adaptability, low O(N) memory demand and their capability to model heterogeneous (isotropic or anisotropic) geometries. These attributes are essential in developing general-purpose codes for electromagnetic analysis/design, including antenna design and characterization. Other attributes of the finite element method are:

- input data to FEM software can be extracted directly from commercially available (i.e. well-supported) solid modeling packages which run on all popular workstation platforms and are well documented. This is particularly important to problems in antenna analysis and design, where a high degree of geometrical fidelity must be maintained(see Figure 1 for examples of antenna and circuit meshes).

- FEM is totally insensitive to the material composition of the radiating or scattering structure. Also, resistive/material and impedance boundary conditions are readily implemented in a modular fashion.

- Being a near-neighbor method, new "physics" can be added to the FEM codes without a need to alter the original code structure. Neither the moment method nor the finite difference method share this feature.

- Being a frequency domain method, the FEM is ideal for antenna analysis and design purposes. Established hybridizations of the FEM with moment method and ray methods provide an added advantage by delivering the most adaptable and efficient code when compared to other approaches.

- Advances in mesh termination schemes have relaxed accuracy compromises with that aspect of the method. Also, the FEM can benefit from recent fast integral equation algorithms. One may therefore think of the FEM as the core method for treating the heterogeneous volumetric structures including feeds and loads

- Fine geometrical details such as those found in spiral antennas and the feeding structure can be modeled without sacrificing geometrical fidelity and CPU efficiency.

- The immediate output of FEM codes is the near zone fields which can be readily visualized (superimposed with the geometry) and further processed using commercially available tools. Moment method codes do not share this important feature. This inherent aspect of the FEM codes allows for extraction and visualization of many different parameters as needed by the user.

Several finite element codes have been developed at the Univ. of Michigan for the analysis and design of printed antenna configurations. Typically, the printed antenna configuration is assumed to be recessed in some metallic or coated platform and the various codes differ in the element used for the tessellation of the antenna, the type of platform assumed in the analysis (planar, cylindrical or doubly curved) and the closure condition employed for terminating the finite element mesh. The following codes are available for antenna radiation and scattering analysis.

FEMA-CYL: This code is specialized to cavity-backed antennas recessed in a metallic cylindrical platform. The finite element-boundary integral method is employed for truncating the FEM mesh and to simplify the geometry I/O, cylindrical edge-based shell-elements are used for the discretization of the cavity and the aperture. As a result, the mesh generation is done automatically and this process does not require interaction with the user. Multiple patches and patch arrays (consisting of individual cavity-backed elements or elements on a single substrate) can be considered but their geometry is restricted by the generated grid. This situation is similar to finite difference codes, where the model's boundary is modified to fit the geometry. Consequently, although this code is user-friendly, it is not appropriate for computing the radiation parameters of circular patches or spirals. However, scattering parameters are not as sensitive to the minor geometrical modifications and therefore FEMA-CYL can still be used for computing the scattering of non-rectangular printed antenna configurations and cavities.

FEMA-CYL has been validated for antenna and scattering applications. For fast execution, it employs the FFT for performing the matrix-vector products in the iterative solver. Both, the biconjugate gradient and QMR solvers are available as options and by virtue of the FFT, the code's computational and memory requirements remain at O(N). The system solution generates the fields within and on the surface of the antenna cavity. For the radiation and scattering pattern calculations only the surface electric fields are used to generate equivalent magnetic currents which are then used in the radiation integral with the appropriate platform Green's function to generate the antenna parameters. For antenna excitation, a probe or a coaxial cable feed model is employed and in the case of scattering the excitation become the aperture fields due to the incoming field. When a probe feed model is used, the input impedance is obtained by dividing the computed field at the probe location with the probe current. When a coaxial

feed is employed, the excitation is introduced by setting the fields coinciding with the edge elements bordering the opening of the coax cable equal $\Delta V/\Delta L$, where ΔV is the given potential between the outer and inner surface of the cable and the $\Delta L=b-a$, where b and a correspond to the outer and inner radius of the cable. The center conductor is modeled by setting to zero the mesh fields at location of the vertical conductor which connects the antenna to the cable.

FEMA-CYL can also be used to include the effect of superstrate materials which may extend over the antenna platform. To avoid use of complicated Green's functions the mesh may be extended a small distance over the cavity's aperture and an absorbing boundary condition is employed for truncating the FEM mesh. With this type of mesh truncation, the entire system is sparse and thus the memory and CPU requirement are again O(N) without a need to make use of the FFT. We have found that placement of the mesh at a distance of 0.3λ away from the coating's surface is sufficient to obtain the radiation pattern with reasonable accuracy. As before, the pattern is computed by introducing equivalent electric and magnetic currents placed at the surface of the dielectric coating and then integrated using the radiation integral.

FEMA-BRICK: This code is a special version of the FEMA-CYL code and is applicable only for cavity-backed antennas recessed in planar platforms. It is also referred to as XBRICK in the community and employs brick elements for the discretization of the cavity volume. Thus, as was also the case with FEMA-CYL, the geometry of the antenna must be modified to fit the grid. Because of the simplicity of the Green's function and the use of the FFT on the boundary integral, the XBRICK is the fastest code among all mentioned here. The code is also provided by Wright Laboratory with a graphical interface to enter the geometry and specify the material layers, loads and printed elements.

FEMA-TETRA: This code employs the same FEM formulation as FEMA-CYL and FEMA-BRICK for computing the antenna and scattering parameters of printed antennas recessed in a metallic platforms. In contrast to FEMA-CYL, this code employs tetrahedral elements for modeling the radiating structure and bricks or shells for modeling the feed. As a result, it incorporates maximum geometrical adaptability and has already been employed for the analysis of antenna+feed configurations beyond the capabilities of moment method codes. The FFT can also be invoked for carrying out the matrix-vector products associated with the boundary integral portion of the system in the iterative, thus maintaining an O(N) memory requirement for the entire system. Its accuracy has already been demonstrated for a variety of antennas, including stacked patches and finite arrays in ground planes, printed and slot spirals and aperture antennas. In the case of aperture fed patch configurations, the mesh generation is divided into the regions above and below the feeding slot. The mesh in each region can employ different elements and this is important in avoiding meshing bottlenecks due to the narrowness of the slot. The mathematical connection between the meshes are then mathematically connected by enforcing electric field continuity across the slot. However, since the mesh elements in each regions are different, this condition is implemented by first relating the fields in the slot

to the potential existing across the slot. Using this procedure, there is no need to have coincidence of the edges bordering the slot and is the main reason which allows use of different elements in the regions above and below the slot aperture. The FEMA-TETRA code is interfaced with commercial solid-modeling packages such as SDRC I-DEAS, Hypermesh and PATRAN to enter the geometry and generate the volume mesh. Using such geometry modeling packages and owed to the adaptability of tetrahedrals, nearly any printed antenna configuration can be modeled and analyzed using FEMA-TETRA. As can be expected, because FEMA-TETRA is not specialized to any antenna configuration all geometrical data must be entered through the solid-modeling package and this may be a time consuming task for inexperienced users.

FEMA-PRISM: To avoid the time consuming task of volume mesh generation without compromising geometrical adaptability, this code provides an attractive compromise by making use of prismatic elements. Prisms have triangular faces at their top and bottom faces and as a result they can model any patch configuration with sufficient geometrical fidelity. Typically, since the substrate/substrate is of constant thickness, once the triangular grid is constructed, the prismatic volume mesh can be generated by growing the elements above and below the surface of the printed antenna. Thus use of prismatic elements reduces the mesh generation task from a three-dimensional to a two-dimensional task. That is, the user needs to only generate the surface mesh in the aperture containing the printed antenna and the volume mesh is then generated automatically by specifying the substrate/substrate thickness, aperture size and information relating to the closure condition. FEMA-PRISM makes use of an ABC or artificial absorbers for truncating the mesh at some distance from the aperture surface. As in the other codes, the radiation and scattering patterns are computed using the aperture fields and the corresponding equivalent magnetic currents in the radiation integral with the free space Green's function as the propagator. Also, since a non-integral closure condition is used for truncating the mesh, the antenna platform can be planar or non-planar (doubly curved). It is, of course, assumed that the generated results will not include effects due to platform interactions and to include these it is more appropriate to interface the FEM codes with high frequency analysis codes[11]

References

1. T. Ozdemir and J.L. Volakis, "Users manual for FEMA-PRISM, " Univ of Michigan Radiation Lab. Techn. Report 031307-6-T, March 1996. 15pp.

2. T. Ozdemir and J. L. Volakis, "Triangular prisms for edge based vector finite element antenna analysis," U-M Radiation Laboratory report 031307-4-T, March 1995, 20 pp

3. J. L. Volakis, A. Chatterjee and L.C. Kempel, "A review of the finite element method for three dimensional scattering," *J.Optical Society of America A*, April 1994, pp. 1422-1433,.

4. J.Gong and J.L. Volakis, A. Woo and H. Wang, "A hybrid finite element method-boundary integral method for the analysis of cavity-backed antennas of arbitrary shape," *IEEE Trans. Antennas Propagat.*, **42**, Sept. 1994, pp. 1233-1242.

5.   J.L. Volakis, J. Gong and T. Ozdemir, "Applications of the FEM to conformal antennas," chapter to appear in book edited by Itoh, Silvester and Pelosi, Wiley, 1996.

6.   J.L. Volakis, J.Gong and A. Alexanian "A finite element boundary integral method for antenna RCS analysis," *Electromagnetics*, **14**, No.1, 1994, pp. 63-85,.

7.   J.M. Jin and J. L. Volakis, "A hybrid finite element method for scattering and radiation from microstrip patch antennas and arrays residing in a cavity," *IEEE Trans. Antennas and Propagat.*, **39**, Nov. 1991, pp. 1598-1604,

8.   L.C. Kempel, J.L. Volakis and R. Sliva, "Radiation by cavity-backed antennas on a circular cylinder," *IEE Proceedings-H, ,* 1995, pp. 233-239

10.   T. Ozdemir, J.L. Volakis, et.al. "Finite element scattering and radiation analysis using prismatic meshes and artificial absorbers for conformal domain truncation," 1996 ACES conference Proceeding, pp. 1174 - 1180.

11.   T. Ozdemir, M.W. Nurnberger, J.L. Volakis, R. Kipp and J. Berrie, "A hybridization of finite element and high frequency methods for pattern prediction of antennas on aircraft structures," *IEEE Antennas and Propagat. Soc. Mag,,* June 1996.

12.   Dana Howell and Bruce Rahn, " Antenna WaveFront Simulator(AWFS)," Wright Laboratory(WL/AAI-1), WPAFB, OH 45433

13.   S.W. Schneider, "On the Analysis of a Stacked Patch Antenna", Wright Laboratory Memorandum.

Connectionist Learning Methods for Reinforcement Learning Tasks: *Does Second-Order Information Help?*

H. Brown Cribbs, III
Graduate Research Assistant
Department of Engineering Science and Mechanics

The University of Alabama
P.O. Box 870278
Tuscaloosa, Alabama 35487-0278

July 1996

Connectionist Learning Methods for Reinforcement Learning Tasks: *Does Second-Order Information Help?*

H. Brown Cribbs, III
Department of Engineering Science and Mechanics
The University of Alabama
Tuscaloosa, Alabama
brown@galab3.mh.ua.edu

Abstract

Artificial neural networks (ANNs) have been shown to be useful function approximators. Reinforcement learning (RL) has used ANNs to approximate Q-functions, value functions, etc. While these methods have shown promise, their slow convergence to an optimal policy is frustrating and is an obstacle for their use in RL. This study looks at RL from a connectionist learning point of view and assesses several ANN learning methods. Specifically, backpropagation, backpropagation with momentum, conjugate gradient, and quasi-Newton (one-step secant) learning methods are compared on a simple RL task. The author hopes to provide insight into the behavior of both, first- and second-order methods in an RL setting.

Keywords: reinforcement learning, conjugate gradient, second-order methods, machine learning, neural networks, backpropagation, gradient descent, Quasi-Newton methods.

Connectionist Learning Methods for Reinforcement Learning Tasks: *Does Second-Order Information Help?*

H. Brown Cribbs, III

1 Introduction

Reinforcement learning (RL) is becoming a popular area in machine learning. Many methods for learning in environments where optimal actions are not known ahead of time have been developed (Sutton, 1987; Watkins, 1989). The popularity of artificial neural networks (ANNs) quite naturally has become one method to predict value functions in RL environments (Lin, 1992; Baird, 1993; Baird, 1995; Smith & Cribbs, 1996).

Inclusion of ANNs in RL problems brings with it a number of problems. Slow convergence in learning an optimal policy is one problem and divergence of prediction values is another problem. This study looks at a the first problem, learning speed, while considering divergence.

The motivation of this study is to assess various second-order methods in an RL setting. A comparison between second-order (conjugate gradient, quasi-Newton) methods and first-order (gradient descent, or backprop) methods is made. The basis for comparison is the number of training cycles (epochs) necessary to learn an optimal policy. From this comparison it is hoped that a basis for accelerating the learning process may be developed for online purposes.

2 Connectionist Learning Methods in Review

Many neural network algorithms have been developed to decrease the number of training cycles needed to reach a desirable level of performance. Momentum and higher order methods, like conjugate gradient methods, have been shown to be effective learning accelerators in supervised learning (Rumelhart et al., 1986; Vogl et al., 1988; van der Smagt, 1994).

The slow rate of convergence in connectionist RL systems make the search for faster learning methods im-

perative. Until now, simple methods, that are readily extensible to online learning have been used. However, these simple methods are slow, so the need to investigate the applicability of *second-order* methods is evident. Second-order ANN training methods have traditionally been developed for *batch mode* training. Practical RL applications use the *incremental (online) mode* of training. This study uses the batch mode to determine whether or not second-order information helps or hinders the learning process in connectionist RL. It is hoped that the information gained by assessing these methods in a batch RL environment will help determine whether development of online second-order methods is warranted.

The following section reviews the basis for modern connectionist RL methods—residual algorithms. After the review of residual algorithms, backpropagation, momentum, conjugate gradient, and quasi-Newton methods are discussed in the context of connectionist RL.

2.1   Residual Algorithm Basis for RL

Most RL systems are based in Bellman's equation (Bellman, 1961),

$$Q(x, u) = \left\langle r + \gamma \max_U \left(Q(x', u') \right) \right\rangle, \tag{1}$$

where $Q(x, u)$ denotes the predicted value of performing action u in state x, r represents the *immediate reward* received for performing action u, and $\gamma \max_U \left(Q(x', u') \right)$ is the discounted maximum reward in the next state, x'.

Baird (1995) points out that the *mean squared Bellman residual*,

$$E_2 = \frac{1}{2} \left(r + \gamma \max_U \left(Q(x', u') \right) - Q(x, u) \right)^2, \tag{2}$$

when used as a basis for connectionist learning, must be differentiated properly. This differentiation must be

done with respect to the weights of the ANN. Since Q-values are output from the ANN, it follows that,

$$\nabla_w E_2 = \left(r + \gamma \max_U \left(Q(x', u') \right) - Q(x, u) \right) \left[\gamma \frac{\partial}{\partial w} \max_U \left(Q(x', u') \right) - \frac{\partial Q(x, u)}{\partial w} \right]. \tag{3}$$

This result reflects the fact that the "desired" value, $r + \gamma \max_U (Q(x', u'))$, is a function of the ANN, as is the "actual" output, $Q(x, u)$.

Previous uses of ANNs in RL considered only the derivative of the actual output, this can be termed the "direct approach" to ANN-RL (Baird, 1995). Neglecting the fact that the desired value changes with subsequent weight updates led ANN, RL applications to erratic behavior because convergence to a minimum is not guaranteed unless both partial derivatives, as in Equation 3, are taken. The difference in derivatives,

$$\left[\gamma \frac{\partial}{\partial w} \max_U \left(Q(x', u') \right) - \frac{\partial Q(x, u)}{\partial w} \right] \tag{4}$$

can be viewed as the *residual gradient in the weight space*. Using this notion, Baird (1995) showed that convergence was guaranteed in a simple residual backpropagation network. The next sections review the backprop method and the more computational second-order (conjugate gradient and one-step secant) methods.

2.2   Backpropagation

One of the most used ANN paradigms is the multi-layer perceptron (MLP) network (Zurada, 1992; Haykin, 1994). This method is often called a "backprop" net, referring to a common method of updating the weights in the MLP net—backpropagation.

"Backprop," is a method that passes input patterns through the network in a feed-forward fashion, rates the output of the network using a desired value, and then calculates the weight updates via a backward pass through the net. In more rigorous terms, the backprop method:

- calculates the error between the network's output and the desired output, via

$$E_2 = \frac{1}{2}(d - o)^2,$$ (5)

where E_2 denotes a *mean-square error measure*, d denotes the desired output of the network, and o is the actual output for the network,

- calculates the gradient of error with respect to the output,

$$\nabla_o E = \frac{\partial E}{\partial o},$$ (6)

- applies the chain rule of calculus to form the gradient with respect to the output layer weights,

$$\nabla_w E = \frac{\partial E}{\partial o} \cdot \frac{\partial o}{\partial w},$$ (7)

- applies the chain rule (again) to find the gradient with respect to the inputs to that layer

$$\nabla_z E = \frac{\partial E}{\partial o} \cdot \frac{\partial o}{\partial z},$$ (8)

where z denotes the input to the output layer neurons,

- then calculates the next layer's update via,

$$\nabla_v E = \underbrace{\frac{\partial E}{\partial o} \cdot \frac{\partial o}{\partial z}}_{\nabla_z E} \cdot \frac{\partial z}{\partial v},$$ (9)

where v denotes the "next" layer's weights.

This method of continuously applying and reapplying the chain rule allows the error to be minimized.

Thus,

$$w(t+1) = w(t) + \Delta w(t) \tag{10}$$

updates the weights where:

$$\Delta w(t) = -\eta \cdot \nabla_w E_2. \tag{11}$$

Note that the w in Equation 11 denotes the current layer of weights being updated. The equivalent *residual gradient update* is,

$$\Delta w(t) = -\eta \cdot \left\{ \left(r + \gamma \max_U (Q(x', u')) - Q(x, u) \right) \left[\gamma \frac{\partial}{\partial w} \max_U (Q(x', u')) - \frac{\partial Q(x, u)}{\partial w} \right] \right\}. \tag{12}$$

2.2.1 Momentum

A simple modification to backpropagation can be made to form weight updates based on a weighted average of the current the gradient, $\nabla_w E_2$, and the last update. The *momentum* method, as it is called, takes the approach of updating weights such that the update constitutes a move to improve network performance based on the current error gradient and the last update (Zurada, 1992; Haykin, 1994). A straight-forward version of momentum update is

$$\Delta w(t) = -\eta \cdot \nabla_w E_2 + \alpha_m \cdot \Delta w(t-1), \tag{13}$$

where $\alpha_m \in [0.2, 0.8]$ (typically) (Zurada, 1992). In many cases this simple modification of standard backprop can accelerate training (Rumelhart et al., 1986; Zurada, 1992).

2.3 Conjugate Gradients

Many improvements to ANNs have come from the area of unconstrained, nonlinear optimization. Conjugate gradient methods are one such class of nonlinear optimizers.

Gradient descent based methods (backprop, etc.), may take a "zig-zag" path to a function's minimum (Haykin, 1994). Conjugate gradient methods avoid this "zig-zag" behavior by using special knowledge about

the current search direction, $\mathbf{p}(i)$, and the current gradient, $\mathbf{g}(i) \triangleq \nabla_w E_2$, to form the next search direction.

For quadratic functions of N degrees of freedom, conjugate gradient has been proven to converge in N steps (Shewchuk, 1994; Gill et al., 1981). For nonquadratic problems the process is iterative and must have a convergence criteria. MLP networks fall into the iterative category.

Using the nonlinear optimization view point, the conjugate gradient method for MLP networks takes its initial search direction as the negative error gradient,

$$\mathbf{p}(0) = -\mathbf{g}(0). \tag{14}$$

Subsequent directions are found by,

$$\mathbf{p}(i+1) = -\mathbf{g}(i+1) + \beta(i)\mathbf{p}(i) \tag{15}$$

where $\beta(i)$ may be determined by the Fletcher-Reeves (1964) formula,

$$\beta(i) = \frac{\mathbf{g}^T(i+1)\mathbf{g}(i+1)}{\mathbf{g}^T(i)\mathbf{g}(i)} \tag{16}$$

which is sufficient for near quadratic problems, or when the search is near the minimum (solution). Another popular formulation for $\beta(i)$ is the Polak-Ribière (1969; 1971) formula,

$$\beta(i) = \frac{\mathbf{g}^T(i+1)\left[\mathbf{g}(i+1) - \mathbf{g}(i)\right]}{\mathbf{g}^T(i)\mathbf{g}(i)} \tag{17}$$

which has been shown to perform better when the search is further from the solution on some problems.

Once a search direction has been determined, via Equation 15, the weight update may be performed by

$$\mathbf{w}(i+1) = \mathbf{w}(i) + \eta(i)\mathbf{p}(i). \tag{18}$$

The learning (update) rate, $\eta(i)$, is determined by a line-search algorithm. The line-search used in this study is a quadratic interpolation scheme (Brent's Method[1]).

2.4   Quasi-Newton Methods, *One-Step Secant*

Nonlinear optimization is a topic that has looked at a multitude of issues. One of these issues is convergence of different methods. The Newton method, which uses second derivatives, has been shown to have the strongest convergence guarantee for nonlinear problems (Murray, 1972). Because of the excessive computational time, $O(N^2)$, involved when using second derivatives, its usage in MLP networks is limited to less than 100 weights (typically).

Quasi-Newton methods—methods that approximate second derivative information—have been used to successfully approximate Newton's method and take "second place" in the convergence guarantee category (Murray, 1972). Battiti (1992) discusses several methods, including what he terms a *one-step secant method* (OSS). This method is a variant of the *one-step, (memory-less) Broyden-Fletcher-Goldfarb-Shanno* (one-step, memory-less BFGS) method. This method, when coupled with *exact line searches*, produces mutually conjugate search directions in $O(N)$ time (Battiti, 1992). In Battiti's words,

> The difference with other forms of the conjugate gradient method is that the performance of the *one-step positive definite secant update* maintains the "safety" properties even when the search is executed in a small number of one dimensional trials.

"Safety" properties refer to the requirement of symmetry and positive definiteness of the Hessian (matrix of partial second-derivatives), which assures descent to a minimum (Gill et al., 1981; Shewchuk, 1994).

OSS calculates its search direction via,

$$\mathbf{p}(n+1) = -\mathbf{g}(n) + A(n)\mathbf{s}(n) + B(n)\mathbf{y}(n) \qquad (19)$$

where: $\mathbf{p}(\cdot)$ and $\mathbf{g}(\cdot)$ are defined as before, $\mathbf{s}(n)$ denotes the last update *step*, and $\mathbf{y}(n)$ is the difference in the

---

[1] see *Numerical Recipes in C* (Press et al., 1992, pp. 402–405).

current gradient vector, $g(n)$, and the previous gradient vector, $g(n-1)$. The update coefficients, $A(n)$ and $B(n)$, are found by the scalar products,

$$A(n) = -\left(1 + \frac{\mathbf{y}^T(n)\mathbf{y}(n)}{\mathbf{s}(n)\mathbf{y}(n)}\right)\frac{\mathbf{s}^T(n)\mathbf{g}(n)}{\mathbf{s}^T(n)\mathbf{y}(n)} + \frac{\mathbf{y}^T(n)\mathbf{g}(n)}{\mathbf{s}^T(n)\mathbf{y}(n)}, \text{ and} \tag{20}$$

$$B(n) = \frac{\mathbf{s}^T(n)\mathbf{g}(n)}{\mathbf{s}^T(n)\mathbf{y}(n)}. \tag{21}$$

Using OSS's search direction, a line search, like a quadratic interpolation scheme, is used to minimize the mean-square error (residual) along the search direction, $\mathbf{p}(n+1)$. The guarantee that the update is in a descent direction comes from the one-step positive definite requirement[2].

3 A Simple RL Test Problem, *The Linear Quadratic Regulator*

Linear quadratic regulators (LQRs) are control systems that have been used to evaluate RL methods (Bradtke, 1993; Baird, 1993). The LQR used here is a simple system that has continuous inputs and two discrete actions. The goal of the system is to "center" itself in the $x = 0.0$ position and remain there; hence, the system must discover a control policy to reach this goal. For this comparison the discount parameter, γ, was set to 0.9 in order to add a strong emphasis on future actions (and rewards).

The input to the LQR is the position, $x \in [-1.0, +1.0]$ (continuous interval). The actions, $U : u \in \{-1, +1\}$ (discrete set), govern the state transition, i.e.,

$$\dot{x} = u, \tag{22}$$

and may be viewed in the position update form,

$$x^{t+1} = x^t + u \cdot \Delta t \tag{23}$$

---

[2]see (Battiti, 1992), (Gill et al., 1981) or (Murray, 1972) for a more complete discussion.

Typical Q–values for Linear Quadratic Regulator

Graph Depicts Expected Reward for Action u
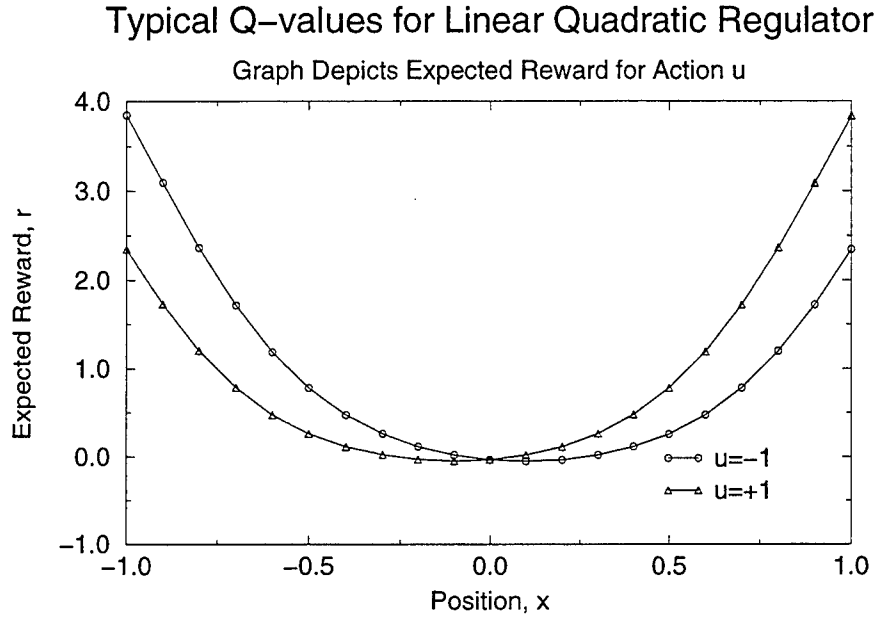
Figure 1: Typical Q-Values for the Linear Quadratic Regulator.

where Δt is the time increment between state transitions.

Using this LQR definition, a simple RL task may be established by defining the reward function to be,

$$r = x^2. \tag{24}$$

As mentioned above, the goal of the LQR is center itself about $x = 0$, thereby, *minimizing* the amount of reward received[3]. The corresponding Bellman equation is:

$$Q(x, u) = \left\langle r + \gamma \min_{U} \left(Q(x', u') \right) \right\rangle \tag{25}$$

where $\langle \rangle$ denotes the expected value of the inside quantity, and $\min_U (Q((x', u'))$ is the minimum Q-value in the action set, U, for the next time step. Figure 1 depicts typical Q-values for the LQR—note that the graph is almost symmetric about $(r = 0.0, x = 0.0)$.

---

[3]Instead of minimizing reward, many texts call this minimizing cost.

3.1 LQR in a Connectionist RL System

The LQR, RL task defined in the previous section was implemented using an ANN for predicting the Q-values. The architecture of the system was arranged such that the inputs to the ANN are:

- the position, x,

- the action, u, and

- a fixed bias, -1.

The output of the ANN was the Q-value, $Q(x, u)$.

This system was tested in *batch learning* mode so that the second-order methods could be evaluated in an appropriate setting. One hundred training samples were formed by dividing the position interval, $X : x \in [-1.0, +1.0]$, into even pieces. Actions for the current time step were generated by pseudo-random "coin flip." The state was then transitioned, via Equation 23, and the minimum Q-value, $Q(x', u')$, was determined. The reward for the transition was determine by Equation 24 and the "desired" value,

$$d = r + \gamma \min_U \left(Q(x', u') \right), \tag{26}$$

was recorded. In order to correctly assess the LQR task under gradient calculations, the previously discussed *residual gradient*[4] was used as a basis for all calculations using the error gradient, i.e., portions of the backprop, conjugate gradient, and OSS methods.

4 Comparison of Connectionist Methods on the LQR, RL Task

The LQR RL task described in the previous section was used to compare convergence time and stability (oscillatory behavior) of the outlined connectionist learning methods. The network consisted of 10 bipolar sigmoid[5] neurons in the hidden layer and a single, linear sum neuron for output. Data points were taken at

---

[4] The residual gradient, Equation 4, was used as the gradient in each method. This means that the network was evaluated for $\nabla_w Q(x, u)$ and $\nabla_w \min_U (Q(x', u'))$.

[5] Bipolar means that the activation function's output ranges from -1 to $+1$.
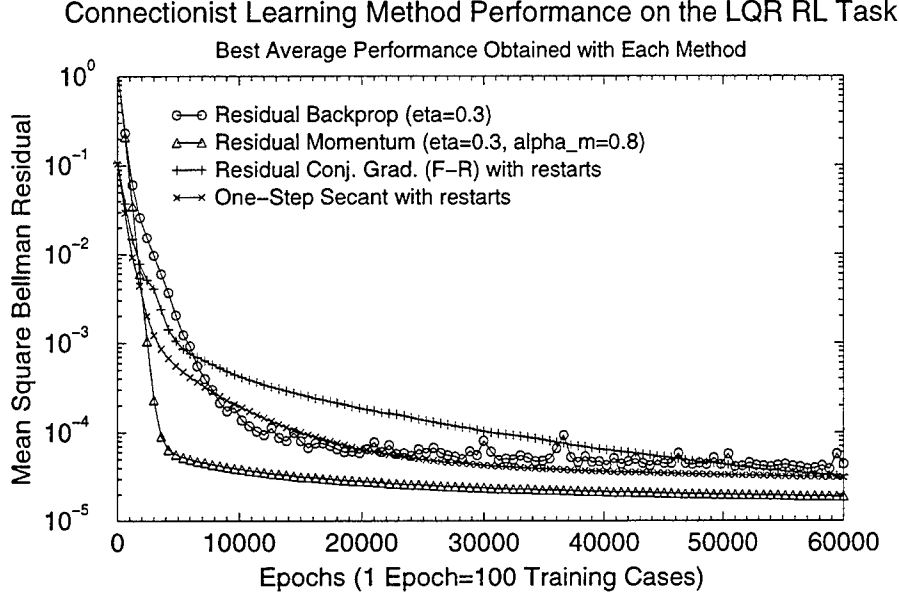
Figure 2: Comparison ANN learning methods for the LQR problem. Each line denotes an average of 10 runs with different random number seeds.

periodically and were exponentially smoothed (continuously averaged). The smoothing constant was fixed so that the comparison could be carefully made on the same scale.

Figure 2 depicts the best, average performance obtained with the learning methods discussed. Each curve reflects the average performance for 10 runs of each method. Backpropagation and momentum methods were evaluated for optimal learning rate and optimal momentum coefficient.

The backpropagation method shows a steady decline, but suffers from the added trouble of finding an optimal learning rate. This limitation caused much time to be spent evaluating learning rates on the interval, $\eta \in [0.1, 0.8]$. The learning rates $\eta \in [0.2, 0.4]$ were shown to be the most suitable with $\eta = 0.4$ displaying oscillatory ("zig-zagging") behavior late in the run (see Figure 3). The learning rate $\eta = 0.3$ provides the best performance; it learns quickly and provides a low error[6] ($\approx 2 \times 10^{-5}$) with little oscillation. The smaller learning rate, $\eta = 0.2$, learns slower with respect to the other two ($\eta = 0.3$ and $\eta = 0.4$) but converges to a similar error level as $\eta = 0.3$ without as much oscillation late in the run. To sum up the backprop method,

---

[6]Note: "error" here means *mean square Bellman residual*
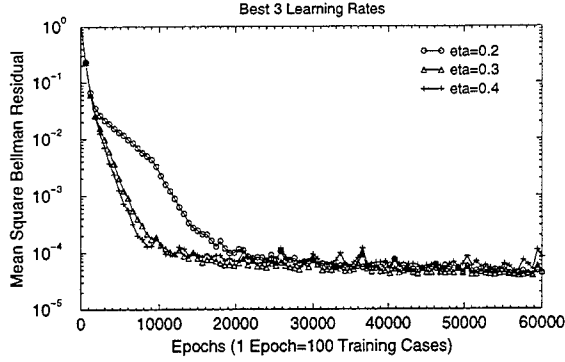
Figure 3: Residual gradient (backprop) average performance for $\eta \in [0.2, 0.4]$.
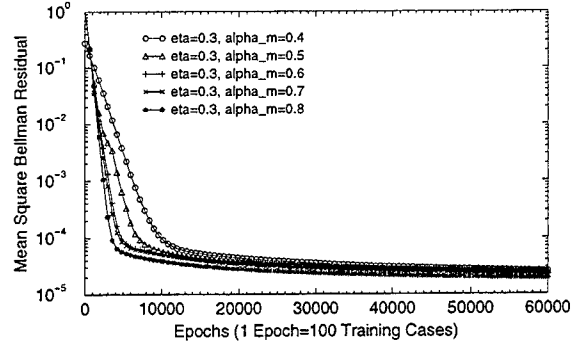
Figure 4: Residual gradient with momentum average performance for $\eta = 0.3$ and $\alpha \in [4, 8]$.

the learning rate $\eta = 0.3$ appears to provide both quick learning speed and stability.

The momentum method showed similar convergence. The higher momentum coefficients, $\alpha_m > 0.5$, appear to provide better performance for the LQR RL problem. Figure 4 shows the best momentum learning curves obtained in testing. Momentum differs from standard residual backprop in that the averaging effects of momentum's update add greater tolerance in selecting a learning rate. Figure 4 also shows that the averaging effects provide little oscillation. It is worthwhile to note that the speed at which momentum converges. Momentum provides the best performance of all the methods ($\eta = 0.3, \alpha_m = 0.8$), which exceeds even the second-order methods with three decimal place accuracy in their line search. Many studies indicate that conjugate gradient is a special case of momentum (van der Smagt, 1994). While momentum did perform the best, the time needed to find optimal settings for both η and α_m are excessive. The second-order methods provide a type of "automatic" parameter setting through their use of a line search.

Figure 2 paints an interesting picture for the second-order methods. The Fletcher-Reeves implementation for conjugate gradient descent reduces the error quickly to around 10^{-2} in about 1000 epochs, but gets stuck for a time (from epochs 2000 to 10000) in a sub-optimal error region. The $\eta = 0.3$ backprop procedure steadily reduced the error, passing conjugate gradient, in the 2000 to 10000 time frame. Conjugate gradient (without restarts) then passes the $\eta = 0.3$ backprop procedure in the low error and stability categories late in the run. This suggests that as the methods reach a near optimum the conjugate gradient method approaches the solution more directly; hence, less error oscillation ("zig-zagging") is evident late in the run.
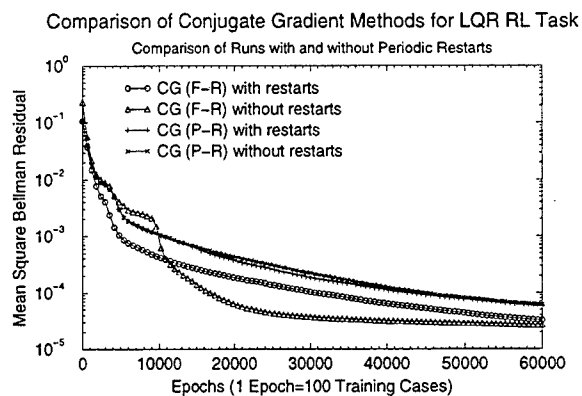
Figure 5: Comparison of conjugate gradient methods with and without periodic restarts.
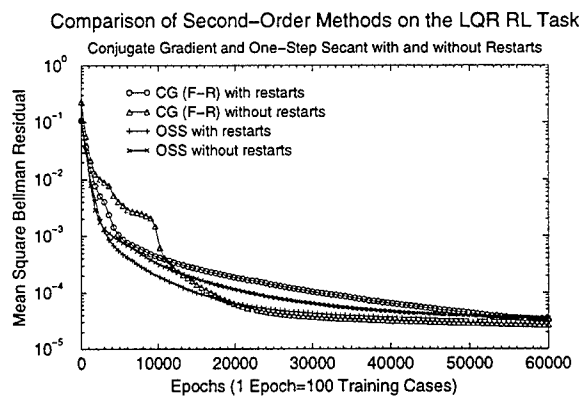
Figure 6: Comparison of conjugate gradient performance versus one-step secant performance. Learning was done with and without periodic restarts for each method.

By adding periodic restarts, the Fletcher-Reeves conjugate gradient method was able to converge slightly faster than the best residual backprop run ($\eta = 0.3$). Many studies suggest that the conjugate gradient method benefits from periodic restarts every N steps, where N, for a ANN, is the total number of weights (Battiti, 1992; van der Smagt, 1994; Haykin, 1994). Restarts increased the performance of the Fletcher-Reeves method in the 2000 to 10000 epochal time frame. Figure 5 shows both conjugate gradient implementations with and without restarts from the negative gradient. Polak-Ribière conjugate gradient did not improve much with periodic restarting. This suggests that the gradient in Fletcher-Reeves drifts slightly, while the Polak-Ribière version provides a seemingly "truer" trajectory. It is likely that the low number of neurons used for the simple LQR is small enough for the differences to be negligible. In larger networks (more connections), periodic restarts may play a more active role in the conjugate gradient procedure.

Battiti's OSS method (1992) shows much promise. It is evident that OSS converges quickly to the 10^{-3} range and then slowly slopes off. The convergence time of OSS is slightly faster (by approximately 2000 to 3000 epochs) than the conjugate gradient methods. With restarts, OSS is better able to slope off at a slightly greater rate (see Figure 6). This implies that the uphill descent directions possible in both Fletcher-Reeves and Polak-Ribière conjugate gradient methods may cause "spiraling" behavior in the descent of the error surface during training. OSS's positive definite Hessian approximation restriction, which is inherent in the BFGS method (Battiti's "safety" properties), may offer better convergence. This follows Murray's

assertion that quasi-Newton methods have stronger convergence properties than conjugate gradient methods for unconstrained optimization problems (Murray, 1972).

5 Conclusion

This study discussed first- and second-order methods for connectionist learning in a simple RL environment. Information gained from these experiments suggest that second-order information is useful and can accelerate the learning process, i.e., less training is needed. This seems to indicate that the residual landscape of LQR is oriented such that even with discontinuous operators, like $\min(\cdot)$, second-order information may be exploited to accelerate learning.

While at the time of writing, the simple momentum method does out perform the second-order methods, the absence of "automatic" settings make momentum a difficult method to implement in many applications requiring adaptation. The automatic setting of the learning rate, η, inherent in the OSS and conjugate gradient methods make for useful application in RL, i.e., simple momentum has to be experimented with to obtain the proper settings for η and α_m.

Clearly quasi-Newton approaches warrant the most immediate attention. OSS shows an advantage in convergence time over conjugate gradient and simple backprop methods. Converging quickly to several orders of magnitude improvement and continuing (albeit slower) to learn without as much "zig-zagging" as in the first-order, residual gradient method makes OSS an interesting area for exploration in RL. OSS represents a simple implementation of the BFGS quasi-Newton approach. Another limited memory BFGS method is that of Nocedal (1980). This limited storage BFGS method (L-BFGS) may provide better performance since it uses multiple steps to update its symmetric, positive definite Hessian approximation.

While this study did not assess RL in its most common (practical) form, incremental (or online) learning, results indicate that pursuit of second-order, incremental learning procedures is warranted. A starting point may be to use a running estimate of the Hessian as suggested by Battiti (1992).

Acknowledgments

References

Baird, L. C. (1993, November). *Advantage updating* (Tech. Rep. WL-TR-93-1146). Wright-Patterson AFB, OH: Wright Laboratory, Avionics Directorate.

Baird, L. C. (1995). Residual algorithms: Reinforcement learning with function approximation. In A. Prieditis, & S. Russell (Eds.), *Machine Learning: Proceedings of the Twelfth International Conference*, San Mateo, CA. Morgan Kaufmann, Inc.

Battiti, R. (1992). First- and second-order methods for learning: Between steepest descent and newton's method. *Neural Computaion* 4, 141–166.

Bellman, R. (1961). *Adaptive control processes: A guided tour.* Princeton, NJ: Princeton University Press.

Bradtke, S. J. (1993). Reinforcement learning applied to linear quadratic regulation. In *Proceedings of the Fifth Conference on Neural Information Processing Systems*, 295–302, San Mateo, CA. Morgan Kaufmann.

Flether, R., & Reeves, C. M. (1964). Function minimization by conjugate gradients. *Computer Journal* 7, 149–154.

Gill, P. E., Murray, W., & Wright, M. H. (1981). *Practical optimization.* San Diego, CA: Academic Press, Inc.

Haykin, S. (1994). *Neural networks: A comprehensive foundation.* Macmillan College Publishing Company, New York: IEEE Press.

Lin, L.-J. (1992). Self-improving reactive agents based on reinforcement learning, planning and teaching. *Machine Learning* 8, 293–321.

Murray, W. (1972). Failue, the causes and cures. In W. Murray (Ed.), *Numerical methods for unconstrained optimization* (107–122). New York: Academic Press, Inc.

Nocedal, J. (1980). Updating Quasi-Newton Matrices with limited storage. *Mathematics of Computation,* *35*(151), 773–782.

Polak, E. (1971). *Computaional methods in optimization.* New York: Academic Press.

Polak, E., & Ribière, G. (1969). Note sur la convergence de methods de directions conjures. *Revue Francaise Information Recherche Operationelle* **16**, 35–43.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1992). *Numerical recipes in c* (Second ed.). New York: Cambridge Univeristy Press.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representation by backpropagating errors. *Nature* **323**, 533–536.

Shewchuk, J. R. (1994). *An introduction to the conjugate gradient method that even an idiot can understand* (Tech. Rep. CMU-CS-94-125). Pittsburgh, PA: School of Computer Science, Carnegie Mellon Uiversity.

Smith, R. E., & Cribbs, H. B. (1996). Cooperative versus competitive system elements in coevolutionary systems. In P. Maes, M. Mataric, J.-A. Meyer, J. Pollack, & S. W. Wilson (Eds.), *From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, Cambridge, MA. The MIT Press/Bradford Books.

Sutton, R. S. (1987). *Learning to predict by the methods of temporal differences* (Tech. Rep. TR87-509.1). Waltham, MA: GTE Laboratories, Inc.

van der Smagt, P. P. (1994). Minimization methods for training feedforward neural networks. *Neural Networks, 7*(1), 1–11.

Vogl, T. P., Mangis, J. K., Rigler, A. K., Zink, W. T., & Alkon, D. L. (1988). Accelerating the convergence of the backpropagation method. *Biologial Cybernetics* **59**, 257–263.

Watkins, J. C. H. (1989). *Learning with delayed rewards.* Unpublished doctoral dissertation, King's College, London.

Zurada, J. M. (1992). *Introduction to artificial neural systems.* St. Paul, MN: West Publishing Company.

CHARACTERISTIC POLYNOMIAL REQUIREMENTS FOR DYNAMIC STABILITY OF RING WING MISSILE CONFIGURATION

Joseph M. DeLong
Graduate Student

Norman Fitz-Coy
Associate Professor
Department of Aerospace Engineering,
Mechanics and Engineering Sciences

University of Florida
Gainesville, FL 32611

CHARACTERISTIC POLYNOMIAL REQUIREMENTS FOR DYNAMIC STABILITY OF RING WING MISSILE CONFIGURATION

Joseph M. DeLong
Graduate Student

Norman Fitz-Coy

Associate Professor

Department of Aerospace Engineering,

Mechanics and Engineering Sciences

Abstract

The stability of a ring wing missile configuration is studied. An inverse symbolic analysis based on implied stability is performed. The linearized equations of motion and the Routh-Hurwitz stability criteria are used to develop the characteristic equations for the stick-fixed equations of motion. A numerical sensitivity analysis on the stability derivatives is performed to identify those derivatives which do not significantly contribute to the stability of the roots of the characteristic polynomial.

CHARACTERISTIC POLYNOMIAL REQUIREMENTS FOR DYNAMIC STABILITY OF RING WING MISSILE CONFIGURATION

Joseph M. DeLong
Graduate Student

Norman Fitz-Coy
Associate Professor
Department of Aerospace Engineering,
Mechanics and Engineering Sciences

Introduction

Although not a new concept, recent interest has developed in the aerodynamics of ring wing missile configurations. Of particular interest is the packaging benefits of a folding ring wing design[2]. The ring wing resembles a cylindrical hoop in its closed position and unfurls to a two dimensional parasol shape under the influence of its stored flexural energy.

(1)   (2)   (3)   (4)
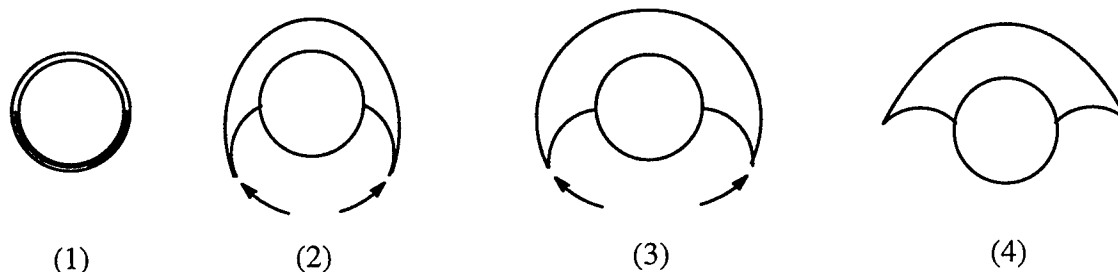
Figure 1. Axial view of ring wing deployment sequence

The ring wing concept enables more compact missile arrangement in an aircraft weapons bay thereby maximizing volume usage. When closed, the ring wing resembles a cylindrical hoop which can be slipped over a generic missile body and fastened at a predetermined position for dynamic stability. These concepts makes the ring wing a more economical and structurally sound method of attaching the primary lift mechanism. This 'wrapped' method of wing storage also makes tube launches an alternative in situations where simple drops are not possible. Figure 2 shows how a generic ring wing missile might look with simple flat plates for tailfins.
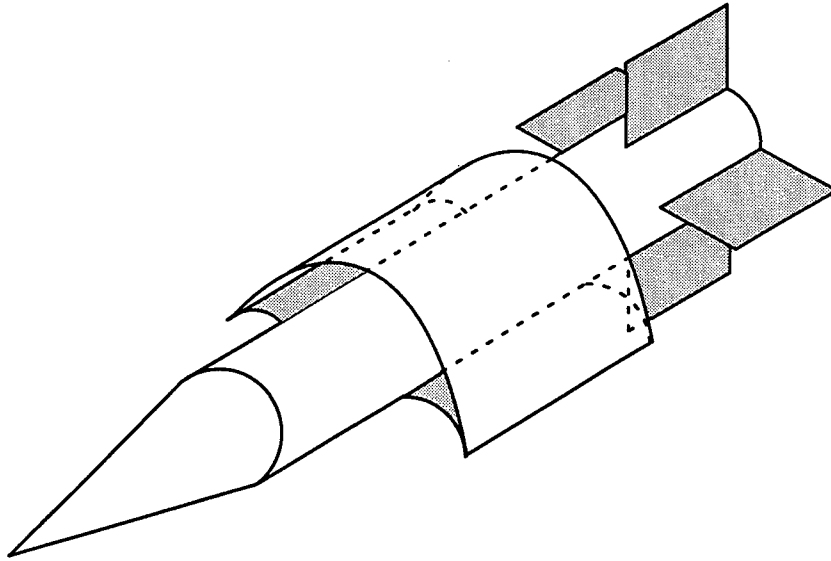
Figure 2. Isometric view of the ring wing missile configuration

Drag is significantly reduced for externally mounted missiles due to a reduction in the wetted area exposed to freestream velocity. Preliminary findings indicate the possibility of more favorable aerodynamics for angles of attack up to twenty degrees compared to those of a conventional cruciform configuration. These two qualities work together to increase the overall lift-to-drag ratio.

In the course of this work, several assumptions have been made to simplify the dynamics to acquire a better understanding of ring wing aerodynamic characteristics. These assumptions as well as the notion of a stable missile configuration will limit the findings herein to that of a qualitative nature. As a first step in the study of ring wing stability, the conclusions drawn will supply further research with valuable insights regarding proper choices for planform and aerodynamic properties.

Problem Statement

The limited knowledge concerning the aerodynamic behavior of a ring wing planform leads to several questions regarding basic design decisions. The bank-to-turn[1] guidance scheme requires the lifting device to remain oriented in a lift producing plane for a majority of the mission as in aircraft concerns. From this requirement comes questions such as roll stabilization

and pitch damping. A longitudinal balance of moments requires some knowledge of the lift force generated by the ring wing.

Also of concern are the vorticity effects on the body and empennage. Design consideration for transitions from subsonic to supersonic regimes are necessary for air-to-ground missiles whose mission requires deployment from supersonic flight conditions. Downwash characteristics affect the efficiency of the tail controls and may vary significantly depending on the flight regime.

Additional information is required to provide insight into the stability characteristics of the ring wing concept. Ordinarily, wind tunnel tests provide aerodynamic coefficients and determine the stability characteristics of the missile. In the absence of wind tunnel data, we can specify desired stability requirements and continue with a stability analysis as if we know the behavior of the ring wing. A stability analysis is then conducted to determine the necessary aerodynamic coefficient ranges for a given geometry configuration and reference flight condition. With an established reference flight condition, the linearized equations of motion can be developed. A graphical interpretation of this study is presented in Figure 3.
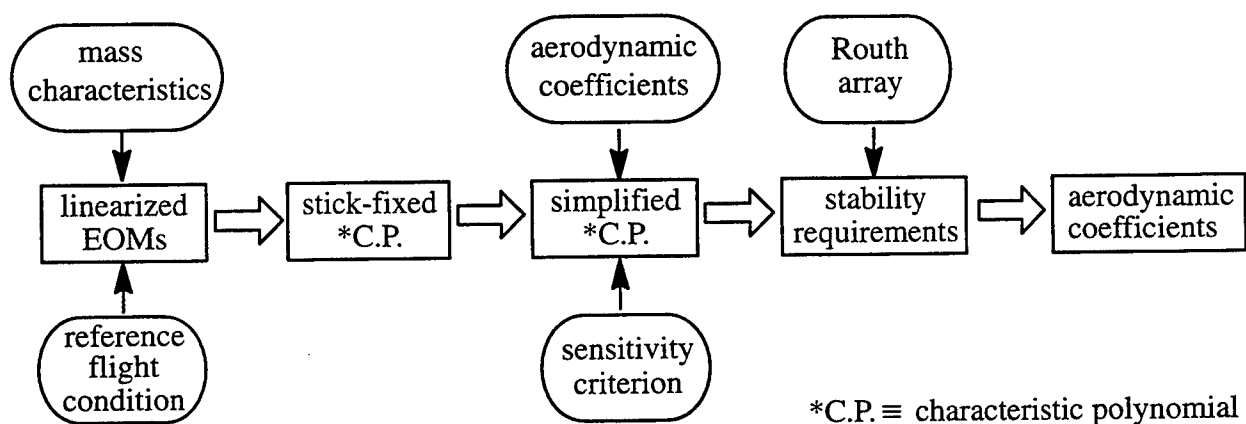
Figure 3. Flowchart indicating the process by which the aerodynamic coefficients are derived

Methodology

This analysis proceeds under the assumption that the missile is stable (i.e., the real parts of the characteristic roots are negative) although modern missile designs incorporate an unstable configuration with computer assisted control to produce a more maneuverable missile. This premise is not overly restrictive since the purpose is to generate insight into the aerodynamics of the ring wing. A location for the center of gravity (cg) is chosen based on experience gathered from similar missile configuration designs with planar wings. Weight and geometry characteristics are arbitrarily chosen and conform to a typical missile configuration design. The geometry and mass characteristics used in this analysis are summarized in Figure 4.

Tail	_Body_	_Wing_
$S_t = 31.80$	$X_{cg} = 13.64$	$X_{LE} = 14.00$
$\bar{c}_t = 5.00$	$L = 34.3$	$X_{cp_w} = 16.4$
$X_{cp_t} = 30.53$	$W = 32\ lb$	$b = 10.28$
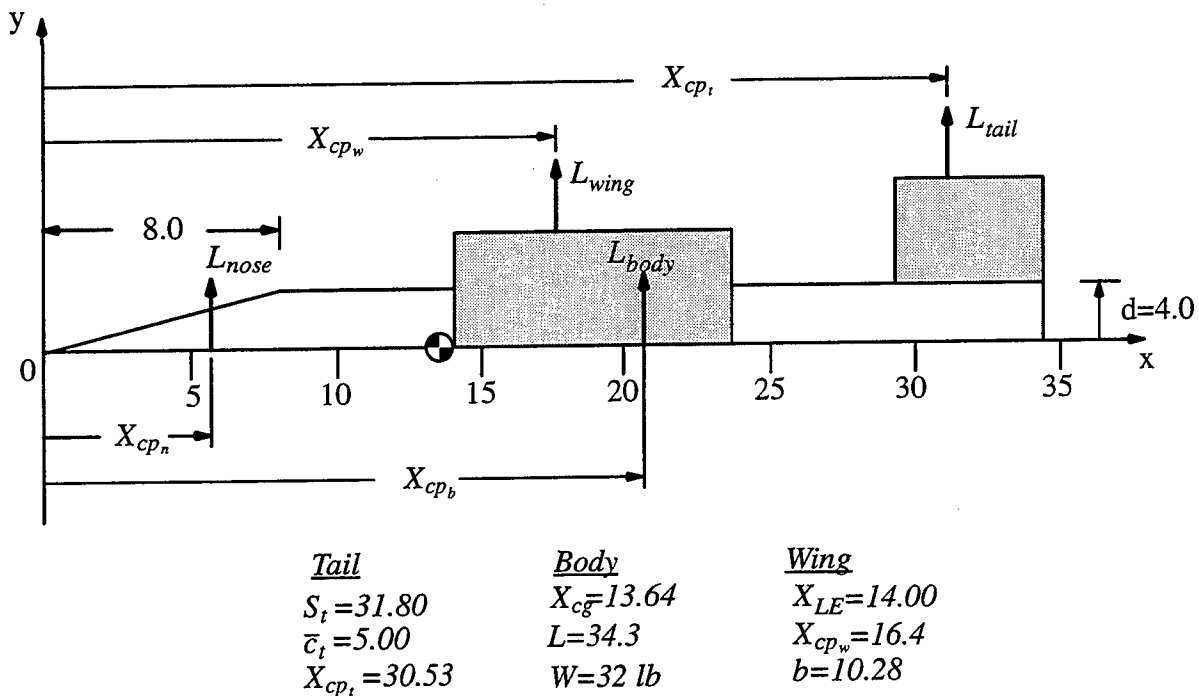
Figure 4. Design model mass and geometry characteristics (all dimensions specified in inches)

A static stability analysis in the design process requires some preliminary information about the missile mass and geometry from which a balance of moments is performed and control surfaces are positioned on the design model. The moments are calculated based on contributions from the nose, wing, body and tail. In this work, influence coefficients for the body's effect on

the wing and tail as well as those for the wing and tail effect on the body are neglected for simplicity. Tail lift coefficients, roll damping and general longitudinal moments of inertia were approximated with the help of PRODAS[4], an aerodynamics prediction code for axisymmetric missiles which combines fluid dynamics theory and a data bank of test results to estimate the aerodynamic coefficients for the proposed model. Due to its axisymmetric modeling restrictions, no information regarding lateral aerodynamics is directly specified by PRODAS. However, by comparing the aerodynamic similarities of sideslip and angle-of-attack, the lateral aerodynamic coefficients were approximated. In the absence of a suitable missile aerodynamics database, the products of inertia were estimated from a table of military fighter aircraft characteristics[3].

A nose cone normal force was estimated to be two per radian based on the equation presented by Chin[5]. A normal force approximation for the body's contribution to lift is also discussed by Chin. Several assumptions have been made in order to define the reference flight condition and make approximations for the aerodynamic coefficients.

1. flat, inertial earth
2. homogeneous atmosphere
3. rectangular, flat, tailfins with zero sweepback
4. unpowered flight
5. rigid, homogeneous missile body
6. missile has constant mass
7. x-z plane is only plane of symmetry
8. lift acts at 25% of ring wing length
9. steady, level flight
10. flight path angle of 15 degrees
11. use of stability axes
12. tail downwash is negligible
13. no wing incidence
14. cg lies on missile axis

The uncontrolled (stick-fixed) dynamic equations of motion are linearized using a small disturbance theorem[8] which states that the motion of the missile consists of small deviations about the steady-state reference flight condition. The motion variables can be written as

$$
\begin{aligned}
u &= u_0 + \Delta u & p &= p_0 + \Delta p \\
v &= v_0 + \Delta v & q &= q_0 + \Delta q \\
w &= w_0 + \Delta w & r &= r_0 + \Delta r
\end{aligned}
\tag{1}
$$

Since we are interested in the dynamic motion of the missile, the perturbations become the variables of interest and the reference condition is eliminated. Since the ring wing is required to be oriented in a lifting plane much like aircraft, it is logical to assume that the missile dynamics are composed of two distinct modes (i.e., longitudinal and lateral). Intentions to implement a bank-to-turn autopilot further support the notion of two modes. The longitudinal mode consists of the motion and force variables (u,w,q,X,Z,M) and the lateral mode consists of (v,p,r,Y,L,N). This permits a reduction of the 6-DOF system of equations into two decoupled 3-DOF systems. The forces and moments are perturbed about the reference flight condition as described in equation (1). For example, the x-force equation is expanded in terms of the longitudinal motion variables as follows:

$$
\Delta X = \frac{\partial X}{\partial u}\Delta u + \frac{\partial X}{\partial w}\Delta w
\tag{2}
$$

where it is assumed that the dependence of ΔX on the other motion variables is negligible. The perturbations of the remaining forces and moments can then be expanded in terms of their associated motion variables. Applying the small disturbance technique to the longitudinal set of equations (X,Z,M) and taking the velocities as the state we have

$$
\begin{bmatrix} \Delta \dot{u} \\ \Delta \dot{w} \\ \Delta \dot{q} \\ \Delta \dot{\theta} \end{bmatrix} =
\begin{bmatrix}
X_u & X_w & 0 & -g\cos\theta_0 \\[2mm]
\dfrac{Z_u}{(1 - Z_{\dot{w}})} & \dfrac{Z_w}{(1 - Z_{\dot{w}})} & \dfrac{(u_0 + Z_q)}{(1 - Z_{\dot{w}})} & \dfrac{-g\sin\theta_0}{(1 - Z_{\dot{w}})} \\[2mm]
M_u + \dfrac{M_{\dot{w}}}{(1 - Z_{\dot{w}})} & M_w + M_{\dot{w}}\dfrac{Z_w}{(1 - Z_{\dot{w}})} & \dfrac{M_q + M_{\dot{w}}(u_0 + Z_q)}{(1 - Z_{\dot{w}})} & \dfrac{-M_{\dot{w}}g\sin\theta_0}{(1 - Z_{\dot{w}})} \\[2mm]
0 & 0 & 1 & 0
\end{bmatrix}
\begin{bmatrix} \Delta u \\ \Delta w \\ \Delta q \\ \Delta\theta \end{bmatrix}
\tag{3}
$$

The stability derivatives (X_u, X_w, Z_u, Z_w, etc.) are defined in the Appendix. The characteristic equation of the stick-fixed equations of motion (3) is a fourth order polynomial of the form

$$As^4 + Bs^3 + Cs^2 + Ds + E = 0 \tag{4}$$

where

$$A = \frac{1}{u_0}(u_0 - Z_{\dot{\alpha}}) \tag{5}$$

$$B = \frac{1}{u_0}[-M_{\dot{\alpha}}(u_0 + Z_q) - (X_u + M_q)(u_0 - Z_{\dot{\alpha}}) - Z_\alpha] \tag{6}$$

$$C = \frac{1}{u_0}\{X_u[M_{\dot{\alpha}}(u_0 + Z_q) + M_q(u_0 - Z_{\dot{\alpha}}) + Z_\alpha] - M_\alpha(u_0 + Z_q) + M_{\dot{\alpha}}g\sin\theta_0 -$$
$$X_\alpha Z_u + Z_\alpha M_q\} \tag{7}$$

$$D = \frac{1}{u_0}\{g\sin\theta_0(M_\alpha - X_u M_{\dot{\alpha}}) + g\cos\theta_0[Z_u M_{\dot{\alpha}} + M_u(u_0 - Z_{\dot{\alpha}})] -$$
$$X_\alpha M_u(u_0 + Z_q) + X_u[M_\alpha(u_0 + Z_q) - Z_\alpha M_q] + Z_u X_\alpha M_q\} \tag{8}$$

$$E = \frac{1}{u_0}\{g\sin\theta_0(X_\alpha M_u - X_u M_\alpha) + g\cos\theta_0(Z_u M_\alpha - Z_\alpha M_u)\} \tag{9}$$

Equation 4 constitutes the stick-fixed characteristic polynomial of Figure 3. $Z_{\dot{\alpha}}$ and Z_q are eliminated immediately upon inspection of equations (5) through (8) due to their insignificance compared to the freestream velocity, u_0. Applying the Routh stability array[7], the following criteria result:

$$A \geq 0 \tag{10}$$
$$B \geq 0 \tag{11}$$
$$(BC - AD) \geq 0 \quad \Rightarrow \quad BC \geq AD \tag{12}$$
$$\frac{D(BC - AD)/B - BE}{(BC - AD)/B} \geq 0 \quad \Rightarrow \quad D(BC - AD) - B^2E > 0 \tag{13}$$
$$E \geq 0 \tag{14}$$

These stability criteria dictate the necessary conditions on the stability derivatives in order to have a stable missile configuration. (Stable here indicates that the real parts of the characteristic roots are negative.) However, equations (12) and (13) become tedious to work

with. A simplified form of the characteristic polynomial coefficients is sought to replace those presented in equations (5) through (9). To this end, we perform a numerical sensitivity analysis on the coefficients to determine which stability derivatives, if any, are negligible.

This analysis requires some knowledge about the stability derivatives for this particular configuration. In an effort to choose reasonable stability derivative estimates, we use the aerodynamic coefficients from the PRODAS model. Again, the axisymmetric restrictions of this code made it necessary to modify the PRODAS results to generate lateral aerodynamic coefficients.
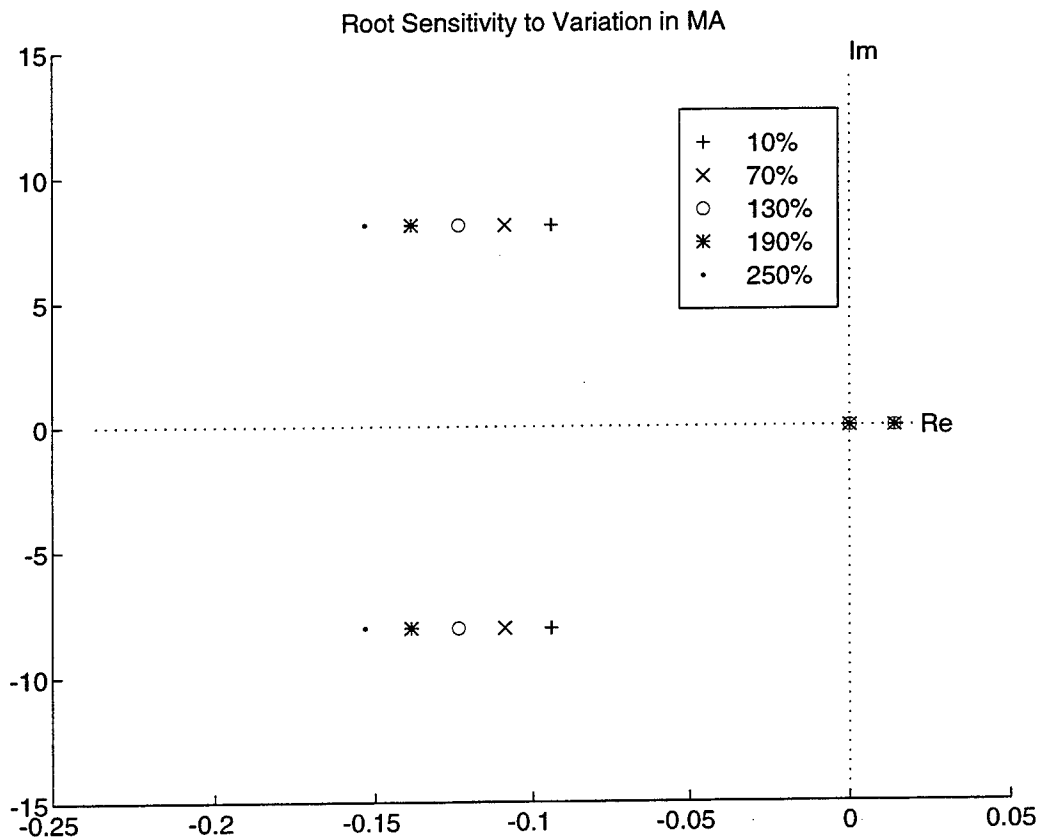
Figure 5. Root migration for $M_{\dot{\alpha}}$ stability derivative

With a set of estimated stability derivatives, a Matlab code was written to plot the characteristic roots on the complex plane. The stability derivatives are then perturbed about the estimates and we trace the root migration paths on the complex plane. We hope to identify those stability derivatives whose roots are not significantly sensitive to variations of the associated stability derivative (i.e., those which do not migrate quickly across the complex plane). Figure 5 is a sample plot from the numerical analysis. $M_{\dot{\alpha}}$ is varied from 10–250% of its value and the characteristic roots of equation (4) are plotted for each variation of this stability derivative. It is easy to see that increasing the magnitude of this stability derivative drives the roots toward a more stable position.

A similar derivation was performed for the lateral set of equations of motion which produced another fourth degree characteristic polynomial. The same numerical analysis was applied to these lateral stability derivatives. The results of the root migration plots for each of the longitudinal and lateral stability derivatives is summarized in Table 1.

Table 1. Summary of root migration plots

Mode	Significant	Insignificant
Longitudinal	Z_α , M_α , $M_{\dot{\alpha}}$, M_q	Z_u , X_u , X_α , M_u
Lateral	L_p , N_β , N_r	Y_β , Y_p , Y_r , L_β , L_r , N_p

Results

The numerical analysis performed on the coefficients of the longitudinal and lateral characteristic polynomials indicated that a majority of the stability derivatives had negligible influence on the root locations of the characteristic polynomials. Two of the derivatives, Y_β and L_β, are of particular interest in typical stability studies and the results of Table 1 lead us to suspect inaccuracies in the estimation of the aerodynamic coefficients. Further study is required to determine the reliability of the PRODAS results and subsequent lateral aerodynamic estimations. An inaccurate estimate of mass characteristics may also be contributing to these results.

Conclusion

A inverse stability analysis was initiated to arrive at the aerodynamic requirements for stability of a ring wing missile configuration for a desired set of characteristic roots. The equations of motion were developed under the notion that the dynamics can be adequately described by a longitudinal and lateral mode. Having arrived at a characteristic equation for both modes, a simplified set of coefficients became necessary to make the symbolic solution of the Routh criteria less complex. A numerical sensitivity analysis proved ineffective in identifying the stability derivatives of negligible influence on the characteristic root locations.

Further information is needed to determine a more accurate estimate of stability derivatives for the root sensitivity analysis. In the case where a new set of estimated stability derivatives produce similar results, the Routh criteria must be reduced using the unsimplified characteristic polynomial coefficients.

As a sequel to this study, a nonlinear analysis of the equations of motion might provide insight into the effects of dynamic cross-coupling between the assumed dynamic modes. A vibration study on the ring wing in its open position might also be useful in determining any harmonic tendencies in the transonic flight regime. Perhaps the efforts to acquire aerodynamic qualities of this unfamiliar configuration will fuel future attempts to more completely characterize its aerodynamic stability.

References

1. Arrow, A.
   Status and Concerns for Bank-to-Turn Control of Tactical Missiles
   Journal of Guidance, March-April 1995

2. August, Henry
   Ring Wing Missile for Compressed Carriage On An Aircraft
   AIAA-93-3656-CP, 1993

3. Blakelock, John H.
   Automatic Control of Aircraft and Missiles
   John Wiley and Sons, 1965

4. Bunnett, John R., Hathaway, Wayne H. and Whyte, Robert H.
   Projectile Design and analysis System (PRODAS-81)
   AFATL-TR-81-43, April 1981

5. Chin, S.S,
   Missile Configuration Design
   McGraw-Hill, 1961

6. Etkin, Bernard and Reid, Lloyd Duff
   Dynamics of Flight: Stability and Control
   John Wiley and Sons, 1996

7. Kuo, Benjamin C.
   Automatic Control Systems
   Prentice-Hall, 1995

8. Nelson, Robert C.
   Flight Stability and Automatic Control
   McGraw-Hill, 1989

9. Nielsen, Jack N.
   Missile Aerodynamics
   McGraw-Hill, 1988

Appendix

α	Angle of attack
m	Body mass
$(\ )_0$	Reference condition
δ_E	Elevator deflection
(p,q,r)	Angular velocity vector
(X,Y,Z)	Body force vector
(u,v,w)	Linear velocity vector
(L,M,N)	Body moment vector
L	Lift
D	Drag
d	Body diameter
q_0	Reference dynamic pressure
S_π	Body reference area (πd^2)
C.P.	Characteristic polynomial

$$M_\alpha = \frac{q_0 S_\pi d}{I_{yy}} C_{M_\alpha}$$

$$M_{\dot{\alpha}} = \frac{q_0 S_\pi d^2}{2 I_{yy} u_0} C_{M_{\dot{\alpha}}}$$

$$M_q = \frac{q_0 S_\pi d^2}{2 I_{yy} u_0} C_{M_q}$$

Longitudinal Stability Coefficients

$$C_{D_u} = \frac{1}{q_0 S_\pi} \frac{\partial D}{\partial(u/u_o)} = \frac{u_0}{q_0 S_\pi}\left(\frac{\partial D}{\partial u}\right)$$

$$C_{D_0} = \left(\frac{D}{q_0 S_\pi}\right)_{\alpha=0}$$

$$C_{D_\alpha} = \frac{1}{q_0 S_\pi}\left(\frac{\partial D}{\partial \alpha}\right)$$

$$C_{L_0} = \left(\frac{L}{q_0 S_\pi}\right)_{\alpha=0}$$

Longitudinal Stability Derivatives

$$X_u = -\frac{q_0 S_\pi}{m u_0}\left(C_{D_u} + 2C_{D_0}\right)$$

$$X_\alpha = -\frac{q_0 S_\pi}{m}\left(C_{D_\alpha} - C_{L_0}\right)$$

$$\left(Note:\ For\ (\ )_\alpha,\ \ \alpha \approx \frac{w}{u_0}\right)$$

$$Z_u = -\frac{q_0 S_\pi}{m u_0}\left(C_{L_u} + 2C_{L_0}\right)$$

$$Z_\alpha = -\frac{q_0 S_\pi}{m}\left(C_{L_\alpha} + C_{D_0}\right)$$

$$M_u = \frac{q_0 S_\pi d}{I_{yy} u_0}\left(C_{M_u} + 2C_{M_0}\right)$$

$$C_{L_u} = \frac{1}{q_0 S_\pi} \frac{\partial L}{\partial(u/u_0)} = \frac{u_0}{q_0 S_\pi}\left(\frac{\partial L}{\partial u}\right)$$

$$C_{L_\alpha} = \frac{1}{q_0 S_\pi}\left(\frac{\partial L}{\partial \alpha}\right)$$

$$C_{M_u} = \frac{1}{q_0 S_\pi d} \frac{\partial M}{\partial(u/u_0)} = \frac{u_0}{q_0 S_\pi d}\left(\frac{\partial M}{\partial u_0}\right)$$

$$C_{M_0} = \left(\frac{M}{q_0 S_\pi d}\right)_{\alpha=0}$$

$$C_{M_\alpha} = \frac{1}{q_0 S_\pi d}\left(\frac{\partial M}{\partial \alpha}\right)$$

$$C_{M_{\dot{\alpha}}} = \frac{2 u_0}{q_0 S_\pi d^2}\left(\frac{\partial M}{\partial \dot{\alpha}}\right)$$

$$C_{M_q} = \frac{2u_0}{q_0 S_\pi d^2}\left(\frac{\partial M}{\partial q}\right) \quad (Blakelock, p19)$$

$$N_\beta = \frac{1}{I_{zz}}\frac{\partial N}{\partial \beta}$$

$$N_p = \frac{1}{I_{zz}}\frac{\partial N}{\partial p}$$

Lateral Stability Derivatives

$$Y_\beta = \frac{1}{m}\frac{\partial Y}{\partial \beta}$$

$$N_r = \frac{1}{I_{zz}}\frac{\partial N}{\partial r}$$

$$Y_p = \frac{1}{m}\frac{\partial Y}{\partial p}$$

Lateral Stability Coefficients

$$Y_r = \frac{1}{m}\frac{\partial Y}{\partial r}$$

$$C_{Y_\beta} = \frac{1}{q_0 S_\pi}\left(\frac{\partial Y}{\partial \beta}\right)$$

$$L_\beta = \frac{1}{I_{xx}}\frac{\partial l}{\partial \beta}$$

$$C_{Y_p} = \frac{1}{q_0 S_\pi}\left(\frac{\partial Y}{\partial p}\right)$$

$$L_p = \frac{1}{I_{xx}}\frac{\partial L}{\partial p}$$

$$C_{Y_r} = \frac{1}{q_0 S_\pi}\left(\frac{\partial Y}{\partial r}\right)$$

$$L_r = \frac{1}{I_{xx}}\frac{\partial L}{\partial r}$$

$$C_{l_\beta} = \frac{1}{q_0 S_\pi d}\left(\frac{\partial l}{\partial \beta}\right)$$

RESEARCH AND DEVELOPMENT OF A
HIGH SPEED HIGH VOLTAGE SEMICONDUCTOR SWITCH

Jorge E. González
Graduate Student
Department of Electrical Engineering

Auburn University
200 Broun Hall
Auburn University, AL 36849-5201

RESEARCH AND DEVELOPMENT OF A
HIGH SPEED HIGH VOLTAGE SEMICONDUCTOR SWITCH

Jorge E. González
Graduate Student
Department of Electrical Engineering
Auburn University

Abstract

Research was done on the development of a high speed, high voltage, Field Effect Transistor (FET). A preliminary fabrication process was developed for the switch which was then modified to meet unique fabrication problems. After testing the switch, further modifications were made to the process. As of this writing, the switch is still in the terminal stages of development and fabrication .

RESEARCH AND DEVELOPMENT OF A
HIGH SPEED HIGH VOLTAGE SEMICONDUCTOR SWITCH

Introduction

High voltage switches capable of operating at high speeds and over a wide range of voltages and energies are used in a variety of applications. High voltage switches capable of fast rise times are commonly used in capacitive discharge units (CDUs). In a CDU, energy stored in a capacitor is coupled through a switch into a low impedance transmission line. Voltages in such systems range from a few volts to thousands of volts.

Of particular interest is the use of small scale capacitor discharges to measure the electrical properties of materials as they are heated from solid through liquid to a gas phase. These vaporized materials are used either as plasma sources for physics experiments, or to propel a thin layer of electrically insulating polyimide for high pressure impact studies.

A limiting element of capacitor discharge units is a fast turn on, low impedance high voltage switch. Several types of switches have been used to drive these systems including triggered spark gaps, dielectric breakdown, and mercury vapor switches. Using a high pressure shock wave in a dielectric to produce a transition from dielectric to conductor has also been used as an efficient single shot switch for capacitor discharges. The purpose of this effort is to develop a solid state semiconductor switch which may be an alternative to the shock-driven single shot switch commonly used for testing.

RESEARCH AND DEVELOPMENT OF A
HIGH SPEED HIGH VOLTAGE SEMICONDUCTOR SWITCH

Methodology

Brief Overview of the Switch:

The MOSFET subsequently discussed is a four layer device consisting of (inner layer to surface layer): patterned silicon, silicon dioxide, aluminum contacts, and a layer of Ti-Ni-Au on top of the aluminum. The wafer used is p-type silicon with phosphorous doped n-wells for the source and drain regions. The oxide thickness is approximately 1 μm. This thick oxide is necessary to avoid breakdown of the oxide at high voltages. Above the SIO_2 is a layer of aluminum for the contacts. The aluminum contacts are coated with Ti-Ni-Au to facilitate soldering to the device. Each die is 400 mil X 400 mil. The large size is necessary to enable the switch to handle large currents and to aid in the interconnection process by minimizing parasitic impedances (this switch will primarily be used with copper strip lines).

The complete fabrication process is outlined below:

The first step in this process is to implant the wafer with phosphorous which will form the n-wells for the source and drain regions. A dose of $2x10^{16}$ was implanted with an energy of 170 keV. A 6 hour drive-in at 1000°C yielded a 3 μm junction depth. The following equation gives a good approximation of the junction depth for this process:

$$N(x,t) = \frac{Q}{\sqrt{\pi Dt}} \exp^{-\left(\frac{x_j^2}{4Dt}\right)}$$

where xj is the junction depth. A cross section of the wafer after implantation is shown in Figure 1.

Phosphorous Implant
P Type Silicon
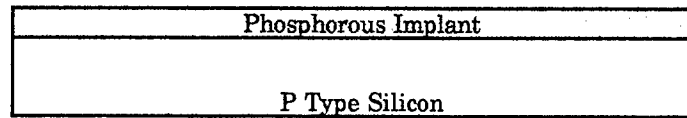
Figure 1

The next step is to deposit aluminum on the surface of the wafer. The aluminum will be used to make continuity measurements across the source and drain regions to ensure that complete etching through of the junction has been achieved. The status of the after aluminum deposition is shown in Figure 2.

Aluminum
Phosphorous Implant
P Type Silicon
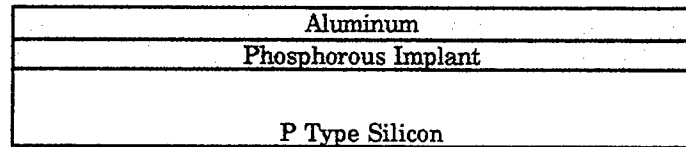
Figure 2

After the aluminum has been deposited an opening must be made in the aluminum so that the silicon can be etched between the source and drain. A thick photoresist is spun on and soft baked at 90°C for 30 minutes to remove solvents and improve adhesion. The status of the wafer at this point is shown in Figure 3.

Photoresist
Aluminum
Phosphorous Implant
P Type Silicon
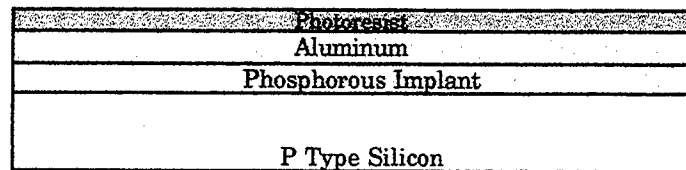
Figure 3

The next step is to align the first mask and then expose and develop the photoresist. This will make an opening in the photoresist at any place where it was not covered by the photomask. A layout of the first mask is shown in Figure 4 and the status of the wafer after developing the photoresist is shown in Figure 5.

Figure 4

Aluminum
Phosphorous Implant
P Type Silicon
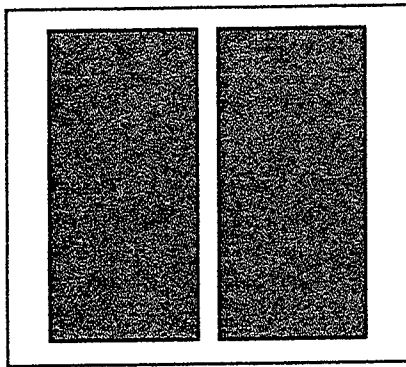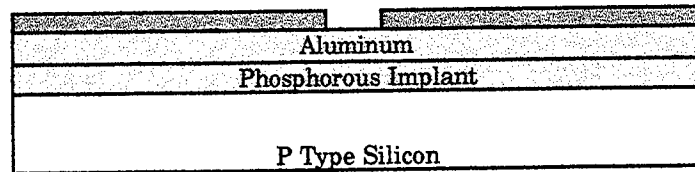
Figure 5

After developing the photoresist, the wafer is dipped in an aluminum etchant (PAE). This will etch the aluminum only where there is no photoresist. Once the aluminum is etched the photoresist can be removed with acetone, methanol, and then a rinse in de-ionized (DI) water. A cross section of the wafer after removing the photoresist is shown in Figure 6.

Aluminum   Aluminum
Phosphorous Implant
P Type Silicon
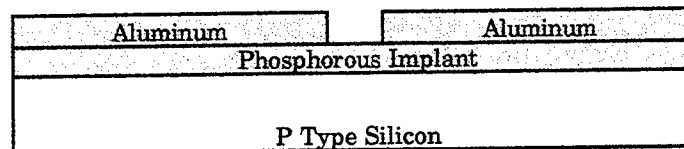
Figure 6

The next step is to isolate the source and drain regions by etching through the phosphorous doped region. This etch was done in a nitrogen tetrafluoride (NF_3) plasma for approximately 330 seconds. An ohmmeter can be used to test for isolation between the source and drain regions. A more conclusive test is to place a voltage across the source and drain regions and then checking for current flow. Negligible current should flow if the n^+ regions are isolated. If the etch is successful, the next step is to remove the aluminum. The method used to remove the aluminum was to dip the wafer in a piranha etch ($7 H_2SO_4 : 3 H_2O_2$). After removing the

9-6

aluminum, the wafer was cleaned and hard baked. The status of the wafer is shown in figure 7.

n+ implant	n+ implant
P Type Silicon	
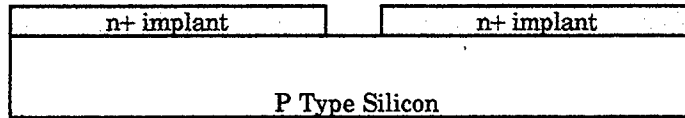
Figure 7

The next step is to grow a thick gate oxide, shown in Figure 8. Wet oxygen at 1000°C for 4 hours yielded an oxide close to a micron thick. This thick oxide is necessary to withstand the high voltages the switch is intended for. Now that the wafer is completely covered with oxide (Figure 9), openings must be etched so the aluminum will contact the n^+ silicon. The layout of the second mask is shown below in Figure 8. The two large openings are the contact holes which will be opened in the oxide. Photoresist is spun on and the second mask is aligned and exposed.

Figure 8

oxide	oxide
n+ implant	n+ implant
P Type Silicon	
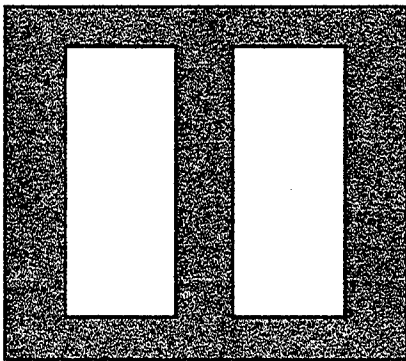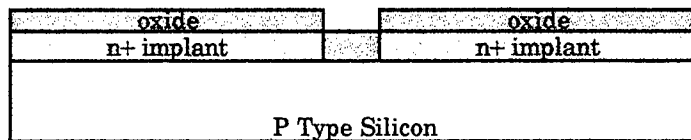
Figure 9

After developing the photoresist, the wafer is dipped in a buffered oxide etchant. This will etch the oxide where the surface is not covered by photoresist. A cross section of the wafer after etching the oxide is shown in Figure 10.
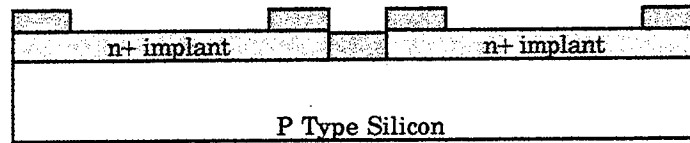
Figure 10

A thorough cleaning of the wafer followed by a dehydration bake should be done before depositing aluminum. An aluminum thickness of 1000 to 2000 Å will sustain the large current through the device. After the surface of the wafer is covered with aluminum, photoresist is spun on and the 3rd mask is aligned and exposed. This mask, shown in Figure 11, will cover the wafer surface where aluminum is desired. After the photoresist is developed, aluminum etchant will be used to remove unwanted aluminum (Figure 12). The aluminum is then annealed to ensure a good contact between the silicon and aluminum.
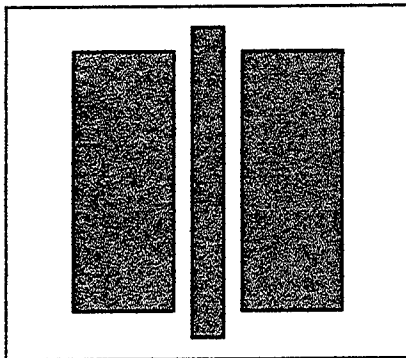
Figure 11
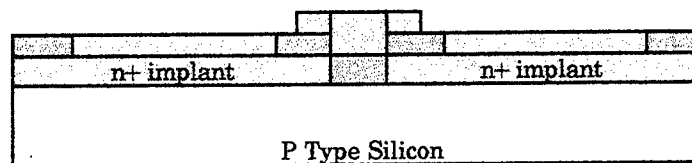
Figure 12

The final step in the process is to deposit Ti-Ni-Au on the aluminum contacts so that the device can be soldered to copper strip lines. Photoresist is spun on and the 4th mask is aligned and exposed. Next, the Ti-Ni-Au is evaporated on the surface and then the photoresist is removed. Any Ti-Ni-Au on the photoresist will be lifted off with the resist. The final cross section of the device is shown in Figure 13.
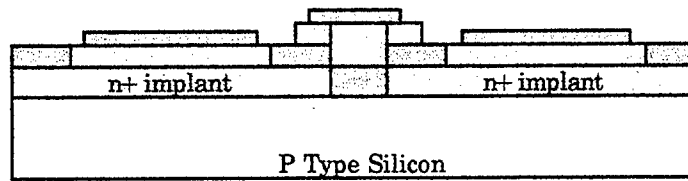
Figure 13

Results

Continuity tests, to check for isolation between the three contacts, indicated a problem with the fabrication process. After the gate oxide was grown, a layer of photoresist was spun on. The photoresist did not completely cover the steep sidewalls between the source and drain, as shown in Figure 14.
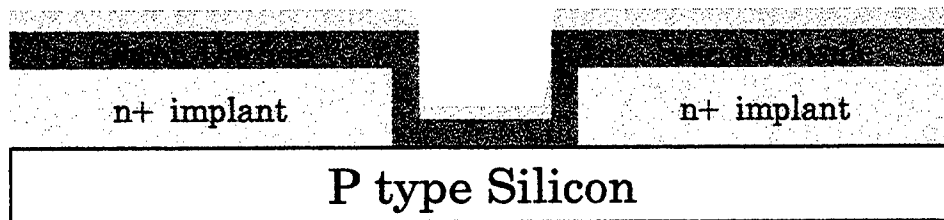
Figure 14

When the wafer was dipped in BOE solution, the oxide on the walls was etched away. The aluminum was then deposited, making contact between the source and drain. This is illustrated in Figure 15.
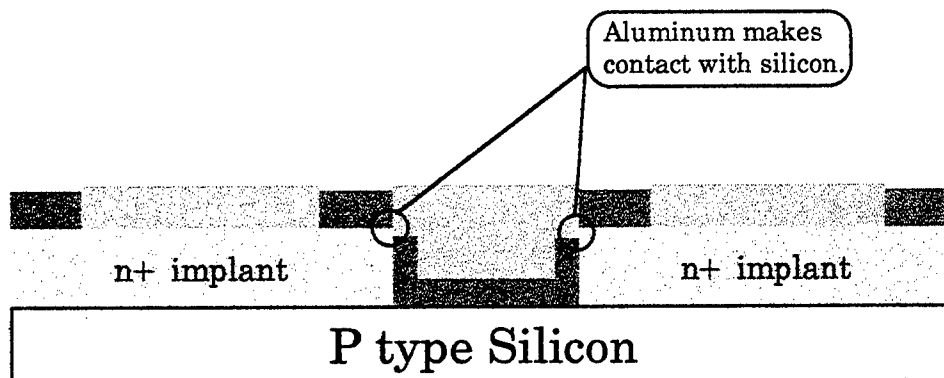
Figure 15

9-9

A crude solution to this problem is to brush on photoresist between the source and drain regions after developing the 2nd mask layer. This will prevent the BOE from etching the vital SiO_2 from the sidewalls. Another solution is to replace the BOE etch with a reactive ion etch. Etching the silicon dioxide in a RIE system will prevent the oxide from being etched away from the steep sidewalls. Both these techniques will be attempted in order to achieve isolation of the source and drain contacts.

INVESTIGATION OF PHOTOLUMINESCENCE INTENSITY SATURATION AND DECAY, AND NONLINEAR OPTICAL DEVICES IN SEMICONDUCTOR STRUCTURES

Jeremy A. Grata
Graduate Student
Department of Physics and Astronomy

Bowling Green State University
104 Overman Hall/Ridge Street
Bowling Green, OH 43403

Final Report for:
Summer Faculty Research Program
Wright Laboratory

September 1996

INVESTIGATION OF PHOTOLUMINESCENCE INTENSITY SATURATION AND DECAY, AND NONLINEAR OPTICAL DEVICES IN SEMICONDUCTOR STRUCTURES

Jeremy A. Grata

Graduate Student

Department of Physics and Astronomy

Bowling Green State University

Abstract

We have observed saturation of photoluminescence peak at low pump intensities in growth-interrupted asymmetric-coupled quantum-well structure. We believe the saturation is due to filling of the exciton states localized at the interface islands. We have observed increase of the photoluminescence decay time as pump intensity increases in the same structure.

Based on our design, a new multilayer structure was grown for demonstrating transversely-pumped counter-propagating optical parametric oscillation and amplification, and achieving surface-emitting sum-frequency generation in a vertical cavity.

We have attempted to mode-lock Ti:Sapphire laser pumped by an Argon laser. We conclude that stability of the Argon laser is crucial for achieving stable mode-locking.

INVESTIGATION OF PHOTOLUMINESCENCE INTENSITY SATURATION AND DECAY, AND NONLINEAR OPTICAL DEVICES IN SEMICONDUCTOR STRUCTURES

Jeremy A. Grata

1. Introduction

Recently, it has been shown that by interrupting sample growth at every interface, one can obtain multiple photoluminescence (PL) peaks with separate emission energies that correspond to the excitonic emissions at interface islands of different sizes [1-4]. Because of the formation of these interface islands, the well widths at these islands generally differ by one monolayer with respect to the designed width in high quality samples [5]. However, the area ratios among all these islands of different well widths are random, which cannot be controlled in growth process. (Without the growth interruption, the recombination of the carriers in the wells with different widths results in the inhomogeneous broadening in the PL spectrum.) At low temperatures, all the carriers generated by the pump will be eventually relaxed down to the lowest energy levels and localized in the islands, resulting in very large carrier densities. If the total area of the islands is small, it would be possible to completely fill exciton states in these islands at relatively low intensities that may manifest as the saturation of the PL peaks. It is worth noting that in growth-interrupted samples, Band-filling effects are spatially-localized effects, due to additional confinement along the interface, similar to situation in quantum dots (i.e. all the islands are spatially isolated).

Recently, surface-emitting green light was obtained [6] by frequency-doubling infrared laser beam ($1.06\,\mu m$) in the waveguide based on periodically modulated second-order susceptibility in alternating $Al_xGa_{1-x}As$ and $Al_yGa_{1-y}As$ ($x \neq y$) layers. When the multilayers are sandwiched between two quarter-wave stacks, large increase in the conversion efficiency was observed [7] though quasi phase-matching was not established. Following Ref. [8], second-order susceptibility of asymmetric-coupled quantum-well (QW) domain structures were measured in the surface-emitting geometry [9]. The maximum conversion efficiency so far is still less than 1%/W. Recently, we proposed a novel practical scheme for implementation of the cascaded nonlinearity using surface-emitting second-harmonic generation (SHG) in the Fabry-Perot cavity. We have shown that such scheme can be efficiently used for optical power limiting and optical phase conjugation at low input power [10]. Most recently [11], we propose to achieve nearly 100% conversion efficiency of SHG for the low input power, by combining quasi phase-matching and cavity enhancements in semiconductor multilayers or

asymmetric QW domain structures. Thus, our investigation leads to the implementation of *practical* frequency doublers which can cover the range from blue to infrared. More importantly, we proposed to implement tunable optical parametric oscillators (OPOs) and amplifiers [12] based on a novel configuration. Frequency doublers, optical parametric oscillators and amplifiers, and the nonlinear optical devices based on the cascaded second-order nonlinearities have potential applications in generation of blue light, generation and amplification of tunable mid-IR light, optical communication, ultrafast detection, sensor protection, real-time holography, or optical lithography.

2. Discussion of Problem

a. Saturation of photoluminescence peak

Previously, we observed [13] saturation of photoluminescence peaks. We believe that it is due to band-filling effects at the interface islands as a result of the growth interruption. The intensities required to observe the saturation reflect the total area of the interface islands, thus the interface roughness.

b. Photoluminescence decay

To determine the characteristic carrier density for completely filling the interface islands, one needs to measure the carrier recombination times. Furthermore, the dependence of the recombination time on the excitation intensity may provide information about nature of the recombination processes.

The sample for studying Sections a and b above was grown by MBE on a semi-insulating GaAs substrate at the temperature of 600 C in collaboration with Naval Research Labs. The epitaxial layers consist of 20 periods, each of which is composed of two narrow asymmetric coupled GaAs quantum wells with the designed thicknesses of 50Å and 65Å, coupled by 35Å-$Al_{0.35}Ga_{0.65}As$ barriers, see Fig. 1. The thicknesses of the barriers between the adjacent periods are 150Å. During the sample growth there is an interruption for 60 seconds at every interface. Because of this growth interruption, interface islands with sizes larger than the average exciton radius are formed, allowing excitons being spatially-localized within these islands with separate optical transition energies from that of free-excitons [1]. As a result, in each designed well the absorption and/or emission peaks are separated from each other corresponding to one monolayer thickness (2.8 Å) difference.

c. Growth of a new structure

200 Å		GaAs	} Cap
1000 Å	X= 0.35	AlGaAs	Layer

Repeat
Period 19X
more
(20 periods
total)

50 Å		GaAs	
35 Å	X= 0.35	AlGaAs	} 1 Period
65 Å		GaAs	
150 Å	X= 0.35	AlGaAs	
5000 Å		AlAs	} Etch Stop
			Layer

Substrate
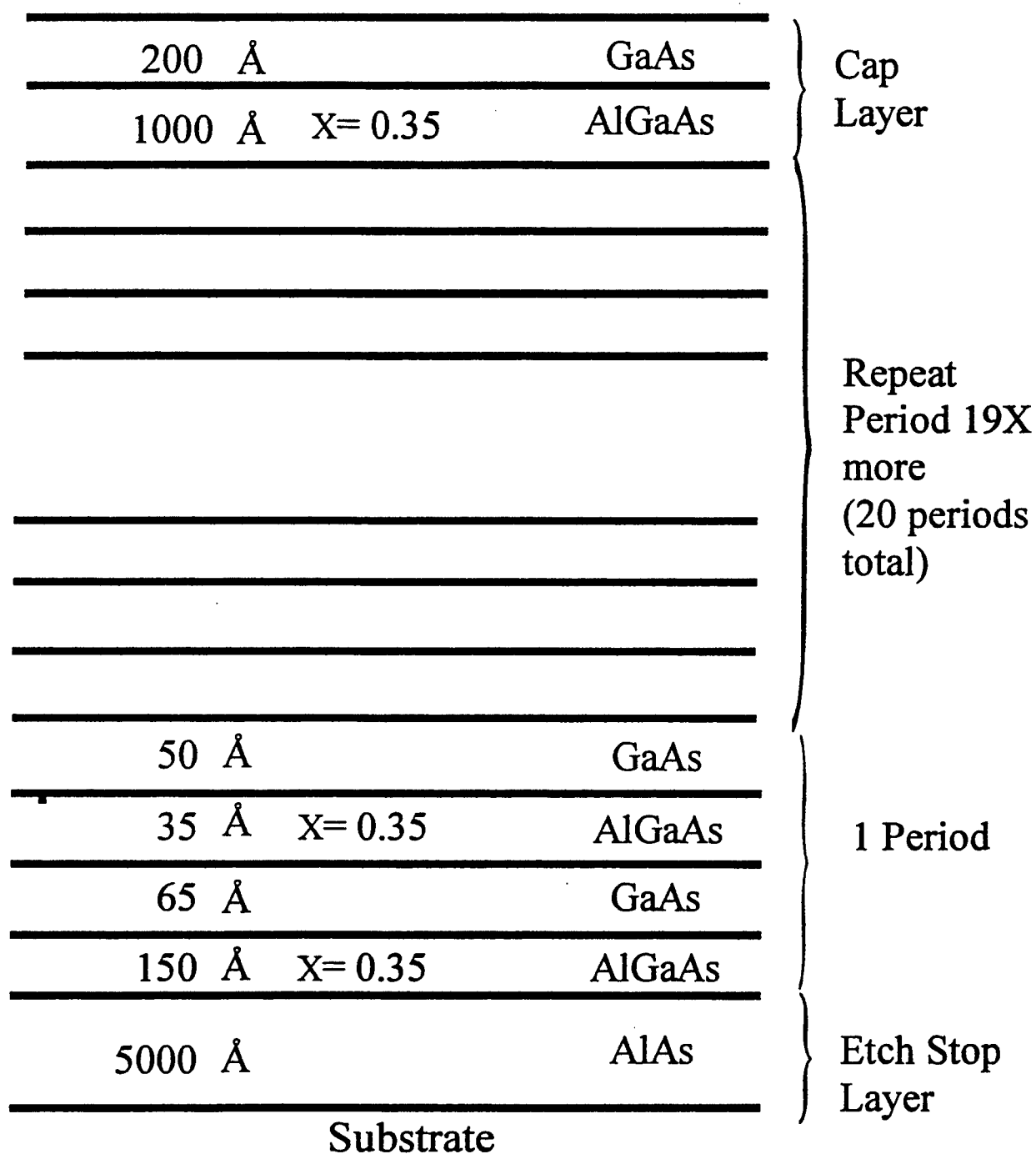
Fig. 1

Previously, we proposed to use semiconductor multilayers to generate surface-emitting second-harmonic radiation [11] to implement transversely-pumped counter-propagating OPOs [11]. Recently, we designed the first structure. The epitaxial layers consist of two Bragg reflectors and alternating layers for achieving quasi-phase matching.

d. Attempt to mode-lock Ti:Sapphire laser

To characterize semiconductor lasers in time-resolved domain, an ultrafast laser source (i.e. mode-locked Ti:Sapphire laser) is required to excite the carriers to the high-energy sub-bands. The relaxation processes can be then probed via different techniques.

3. Methodology

a. Photoluminescence spectrum

Our asymmetric-coupled quantum-well structure is pumped by a CW Argon laser at the wavelength of 5145 Å. The photoluminescence was collected by a monochromator via a lens.

b. Photoluminescence decay

For the measurement of the PL decay, we used a mode-locked Ar^+ laser as our excitation pulse with the pulse duration of 150 ps and output wavelength of 5145 Å. The temporal traces of the PL signal were taken via a streak camera with a time resolution of 20 ps. Fig. 2 shows our schematic set-up.

c. Design of a new structure

Our optimized design of the multilayer structure is based on our vigorous consideration of surface-emitting frequency doublers [Fig. 3(a)] and transversely-pumped counter-propagating OPOs and OPAs [Fig. 3(b)], see Refs. [11,12].

d. Attempt to mode-lock Ti:Sapphire laser

The schematic set-up for Argon-laser-pumped mode-locked Ti:Sapphire laser and the Ti:Sapphire laser cavity are shown in Fig. 4. We have followed the manual for Ti:Sapphire provided by Clark-MXR Inc. The mode structure and stability were determined by eyes after expanding the laser beam. The output laser pulse from mode-locked Ti:Sapphire laser can be sent to Coumarin 460 for generating two-photon fluorescence.

4. Results

Figure 2

(a)   (b)
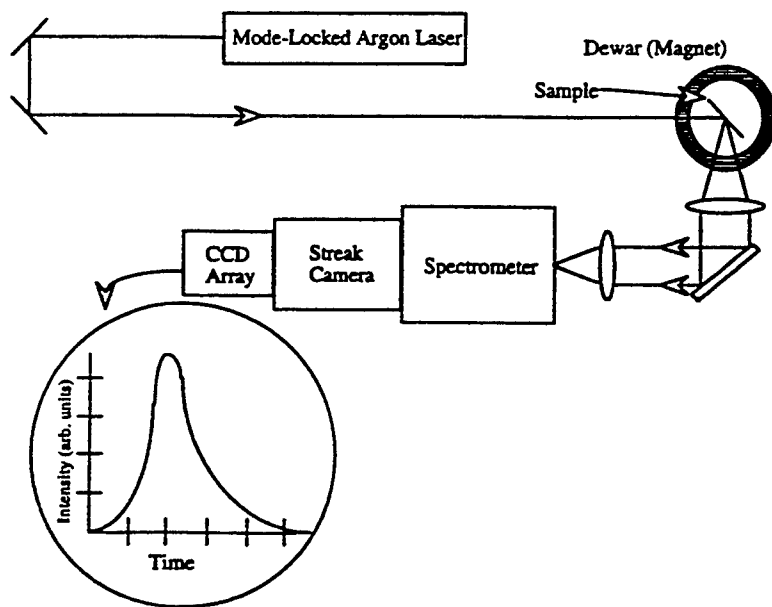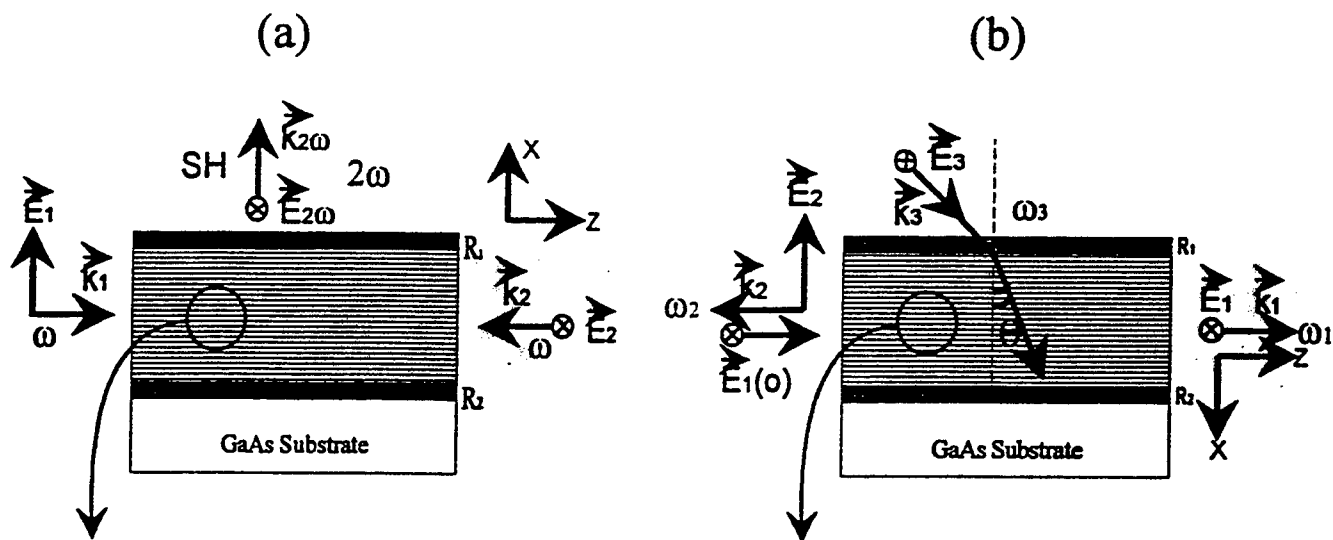
500 Å	GaAs
550 Å	$Al_{0.5}Ga_{0.5}As$
620 Å	AlAs
1080Å	$Al_{0.4}Ga_{0.6}As$
1240Å	AlAs
550 Å	$Al_{0.5}Ga_{0.5}As$
620 Å	AlAs

× 15

× 4

× 15

(100) GaAs Substrate

(c)

Figure 3
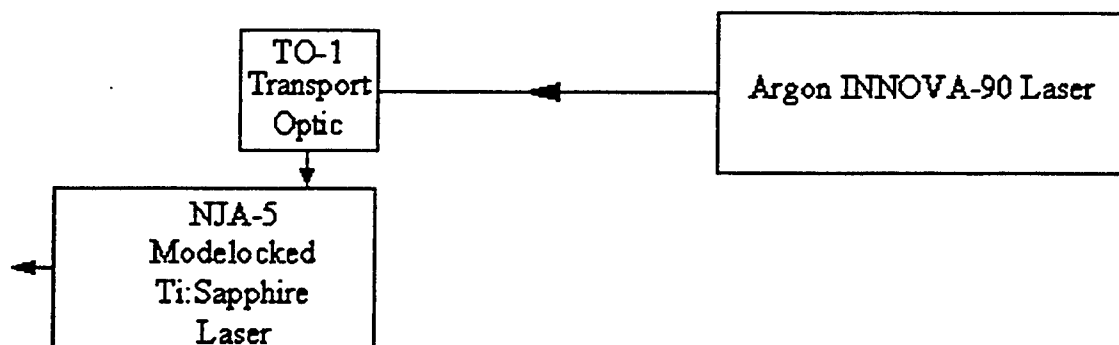
(a)
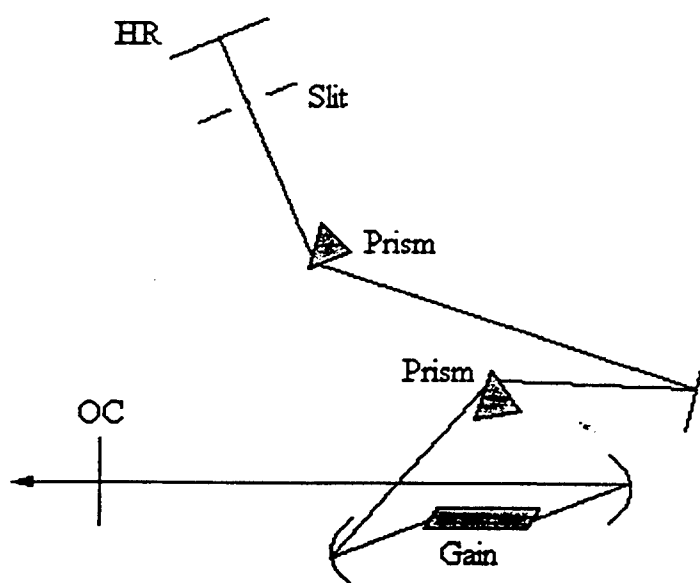
(b) Ti:Sapphire Laser Cavity

Figure 4

a. Saturation of photoluminescence peak

The PL spectra for several pump intensities are shown in Fig. 5. At laser intensity of 9.7 mW/cm^2 there are two emission peaks: the one on the long wavelength side (~ 7780 Å) corresponds to the emission of excitons at the interface islands while the other one (~ 7773 Å) corresponds to the free excitons. When we change the laser intensity from 9.7 mW/cm^2 to 1.4 W/cm^2 at 4 K, we can see that the emission peak for the localized excitons loses its relative strength.

Due to growth interruption, a single PL peak breaks into two because of the formation of interface islands with the size larger than the exciton radius. At low temperatures, all the carriers generated by the pump laser will be eventually relaxed down to the lowest energy levels and localized in the islands, resulting in large carrier densities. If the total area of the islands is small, it would be possible to completely fill exciton states in these islands at relatively low intensities, which manifests as the saturation of the PL peaks. This type of the band-filling effect only occurs at the spatially-localized islands. The laser intensity required to almost completely fill the localized exciton states is more than two orders of magnitude lower than that obtained before [13].

b. Photoluminescence decay

We have made the time-resolved PL measurements in our sample. Fig. 6 shows the typical temporal PL traces detected at the center wavelength of $e_1 hh_1$ emission peak (the excitonic emission peak) as a result of the carrier recombination at the interface islands. At the low excitation intensity (207 W/cm^2), the PL signal at the $e_1 hh_1$ (II) emission peak has a decay time of about 269 ps. As the intensity increases, the decay time increases. As shown in Fig. 6, when the laser intensity is 414, 621, and 828 W/cm^2, the decay time is about 326, 537, and 666 ps.

In the quasi-CW regime, the density of excitons can be determined as

$$N_{cx} = \tau \, \frac{I_{laser} \alpha}{\hbar \omega_{laser}} \tag{1}$$

where I_{laser} is the laser intensity, α is the absorption coefficient, and $\hbar \omega_{laser}$ is the energy of a single photon. The intensity required to completely fill the $e_1 hh_1$ exciton states is about 1.4 W/cm^2. Assuming $\alpha \approx 10000$ cm^{-1} at the pumping wavelength in our experiments, the exciton density is then estimated to be 9.75×10^{12} cm^{-3}. The corresponding area density is 1.46×10^7 cm^{-2}. This carrier density is more than two orders of magnitude lower than that in Ref. [13].

CW PL Plot for
1276.(176, 180, 181, 187)

9.7 mW/cm^2

36 mW/cm^2

0.12 W/cm^2

1.4 W/cm^2

——— 1276.176
········· 1276.180
——— 1276.181
—·—·— 1276.187

Relative PL Intensity

Wavelength (Å)

3500 3000 2500 2000 1500 1000 500 0

7700 7710 7719 7729 7738 7748 7758 7761 7767 7777 7786 7796 7806 7815 7825 7835 7844 7854
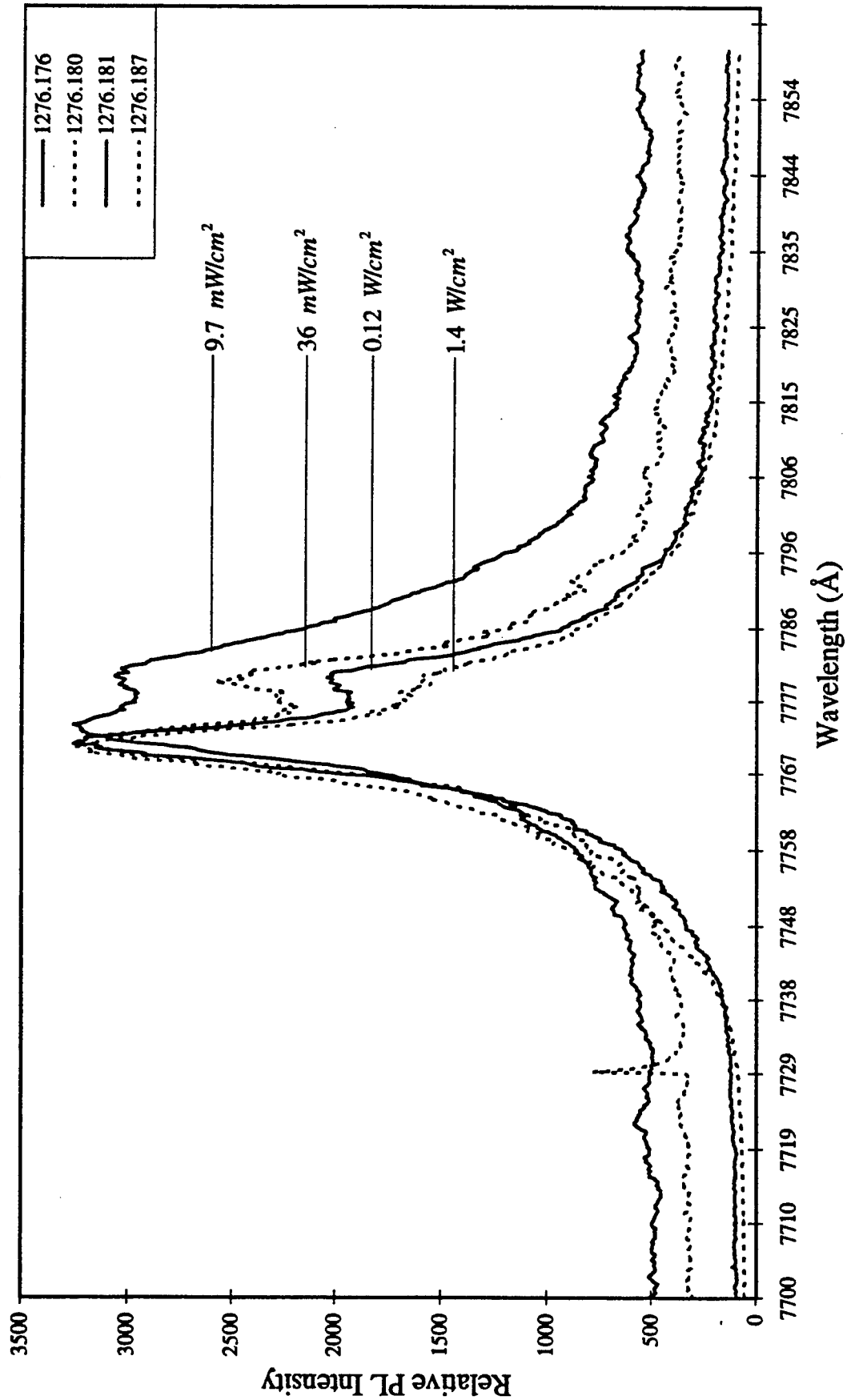
Figure 5

TR PLPlot for
1276.27 1276.54 1276.83 1276.110
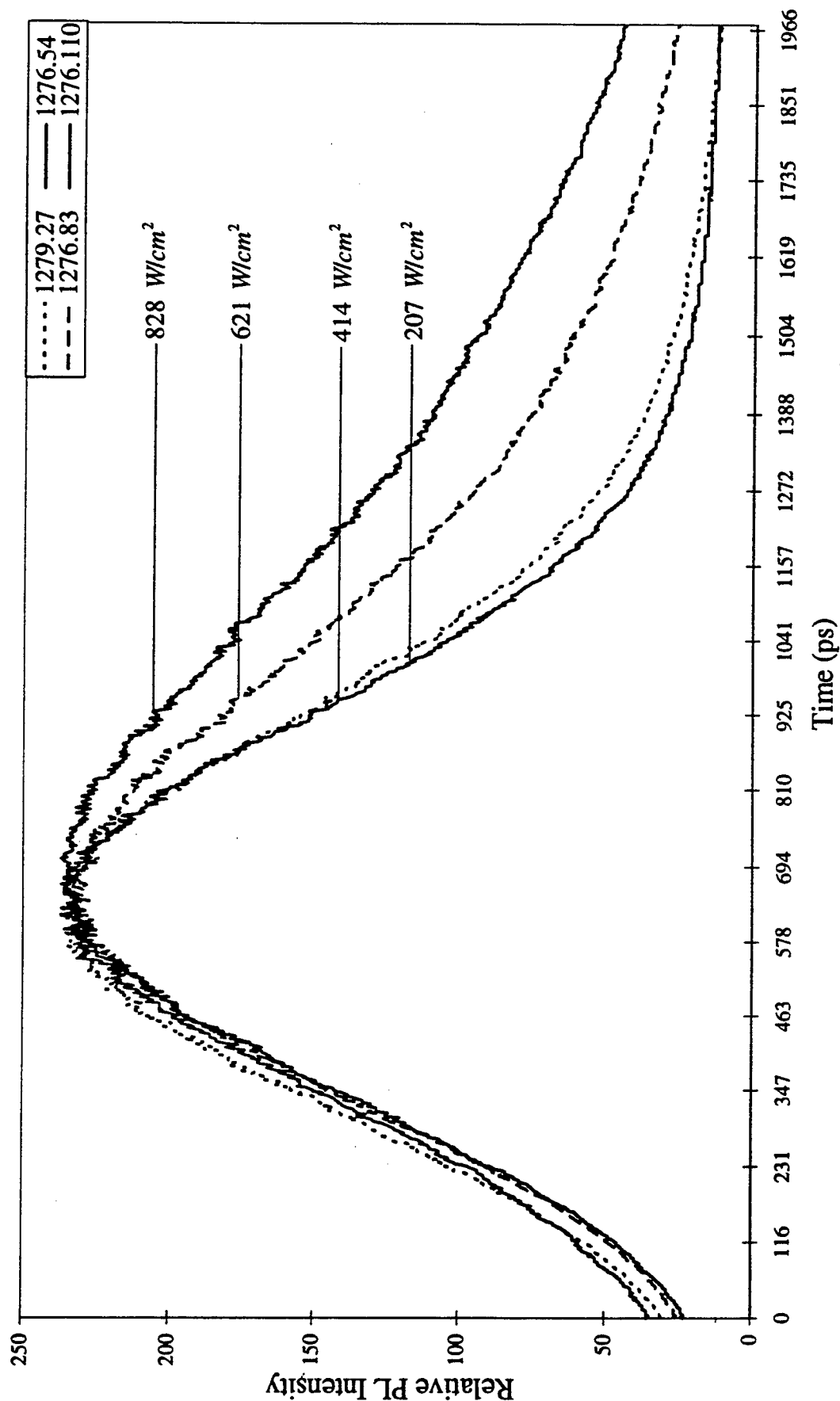
Figure 6

c. Growth of a multilayer structure

Based on Refs. [11,12], we had already designed an optimized structure [see Fig. 3(c)]. We had grown this structure in collaboration with Drs. J. L. Loehr and J. Ehret at Wright Labs. We will test the performance of this structure as an efficient frequency doubler and optical parametric oscillator and amplifier at Bowling Green State University. The pump, input and output wavelengths are designed to be 1.06 μm, 1.58 μm, and 3.23 μm. By changing the incident angle, one can tune the output wavelengths [12].

d. Attempt to mode-lock Ti:Sapphire laser

The Ti:Sapphire laser is pumped by a re-furnished Argon laser (Coherent Innova 90), see Fig. 4. For 4-watt pump power of a multi-line Argon laser, conversion efficiency as high as 20% was achieved in CW regime. We had tried our best to mode-lock Ti:Sapphire laser. We observed two-photon fluorescence in Coumarin 460, however, was not able to stablize the mode-locked laser output. We believe that mode structure and pointing stability could be the problems for mode-locking the Ti:Sapphire laser. When there is no aperture for the Argon laser, high-order modes other than TEM_{00} exist in the Argon output beam profile. In addition, we have crudely estimated the pointing stability as ~ 100 μrad, which is an order of magnitude larger than that required for stable mode-locking.

5. Conclusion

We have observed saturation of photoluminescence peak. We have measured time-resolved photoluminescence decay in growth-interrupted asymmetric-coupled quantum wells. We have subsequently determined decay times and characteristic carrier densities for the observed photoluminescence intensity saturation. We have grown a multilayer structure that can be used to implement an optical parametric oscillator and amplifier and frequency doubler in a novel configuration. Finally, we have tried to mode-lock Ti:Sapphire laser and summarized potential problems causing unstable mode-locked output.

Two journal papers based on our results will be submitted for publication.

6. Acknowledgement

7. References

[1] D. Bimberg, D. Mars, J. N. Miller, R. Bauer, O. Oertel, and J. Christen, Supperlatt. Microstruc. 3, 79 (1987).

[2] M. Kohl, D. Heitmann, S. Tarucha, K. Leo, and K. Ploog, Phys. Rev. B39, 7736 (1989).

[3] D. C. Reynolds, K. K. Bajaj, C. W. Litton, P. W. Yu, J. Singh, W. T. Masselink, R. Fischer, and H. Morkoc, Appl. Phys. Lett. 46, 51 (1985).

[4] K. Fujiwara, K. Kanamoto, and N. Tsukada, Phys. Rev. B40, 9698 (1989).

[5] B. Devcaud, A. Chomctte, N. Roy, B. Scrmagc, and D. S. Katzer, Sur. Sci. 267, 199 (1992); D. Gammon, B. V. Shanabrook, and D. S. Katzer, Appl. Phys. Lett. 57, 2710 (1990).

[6] R. Normandin, R. L. Williams, and F. Chatenoud, Electr. Lett. 26, 2088 (1990); R. Normandin, H. Dai, S. Janz, A. Delage, J. Brown, and F. Chatenoud, Appl. Phys. Lett. 62, 118 (1993); D. Vakhshoori, R. J. Fischer, M. Hong, D. L. Sivco, G. J. Zydzik, G. N. S. Chu, and A. Y. Cho, Appl. Phys. Lett. 59, 896 (1991).

[7] R. Lodenkamper, M. L. Bortz, M. M. Fejer, K. Bacher, and J. S. Harris, Jr., Opt. lett. 18, 1798 (1993).

[8] J. Khurgin, Appl. Phys. Lett. 21, 2100 (1987); Phys. Rev. B38, 4056 (1988); J. Appl. Phys. 64, 5026 (1988); J. Opt. Soc. Amer. B 6, 1673 (1989).

[9] S. Janz, F. Chatenoud, and R. Normandin, Opt. Lett. 19, 622 (1994).

[10] J. B. Khurgin and Y. J. Ding, Opt. Lett. 19, 1066 (1994).

[11] Y. J. Ding, S. J. Lee, and J. B. Khurgin, J. Opt. Soc. Am. B 12, 1586 (1995).

[12] Y. J. Ding, S. J. Lee, and J. B. Khurgin, Phys. Rev. Lett. 75, 429 (1995); Y. J. Ding, S. J. Lee, and J. B. Khurgin, IEEE J. Quant. Electr. 31, 1648 (1995).

[13] A. G. Cui, Y. J. Ding, S. J. Lee, J. V. D. Veliadis, J. B. Khurgin, S. Li, D. C. Reynolds, and J. Grata, J. Opt. Soc. Am. B 13, 536 (1996).

ATMOSPHERIC ATTENUATION MODELING FOR LPI COMMUNICATION PERFORMANCE ANALYSIS

Andrew J. Harris
Graduate Student
Department of Computer Science

Northern Illinois University
2048 Normal Road
DeKalb IL, 60015

ATMOSPHERIC ATTENUATION MODELING FOR LPI COMMUNICATION PERFORMANCE ANALYSIS

Andrew J. Harris
Graduate Student
Department of Computer Science
Northern Illinois University

Abstract

The Low Probability of Intercept Signal Detectability Analysis (LPISDA) computer simulation model is used to determine the vulnerability of communication systems to interception. Currently LPISDA includes only free-space path loss in it's link loss computations. In order for LPISDA to become a useful tool for determining signal vulnerability in all types of weather conditions, it is necessary to find an appropriate weather attenuation model to interface with LPISDA. The goal of this project is to locate the best model to use and integrate into LPISDA.

ATMOSPHERIC ATTENUATION MODELING FOR LPI COMMUNICATION PERFORMANCE ANALYSIS

Andrew J. Harris

Introduction

Low Probability of Intercept Signal Detectability Analysis (LPISDA) is a computer simulation model which evaluates the relative performance of various types of Low Probability of Intercept (LPI) communication systems. It was originally written by Scott P. Francis[1] and later improved upon by Paul W. Purdon. LPISDA was designed for use as an independent tool to analyze the signal detectability of LPI communication systems. The source language is C and LPISDA is compiled to run under DOS on any 286 or higher PC. During execution the user is asked to input[2] parameters which define the communication waveform, and system characteristics. The calculations are performed and the program outputs a "Detectability Matrix" containing an LPI quality factor (Q_{lpi}) for each of 22 different types of intercept receivers. These 22 intercept receivers are partitioned into 4 families (i.e. *Single Filter Radiometers, Filter Bank Radiometers, Scanning Radiometers and Feature Detectors*).

Description of Problem

Modern communication systems must function in all types of weather. The atmosphere contains particulate matter such as water droplets/vapor, dust and various gases. When the size of these suspended particles becomes approximately equal to the wavelength used for communication, some of the wave's energy will be scattered by these suspended particles, causing the signal to be attenuated. LPISDA does not currently model these particulates, and in order for LPISDA to become an effective all-weather system model, it will need to contain a weather simulation.

---

[1]Scott P. Francis; University of Kansas.
[2] Or load from disk.

Many atmospheric transmissivity simulation packages are designed for use with the higher frequency electro/optical bands. Adequate propagation simulations for lower frequencies are difficult to find. Nonetheless, once a suitable model is found, the integration of the transmissivity model into LPISDA can be accomplished. In order for the simulation to be totally integrated, the model's code would need to be written in C and absorbed into the current LPISDA design.

Methodology

In LPI communication systems, an *Overall LPI Quality factor for Air-to-Air and Air-to-Ground Communications* can be defined as follows [1]:

$$Q_{lpi} = 20\log\left(\frac{R_c}{R_i}\right) = 20\log\frac{\left(\dfrac{4\pi R_c}{\lambda_c}\right)}{\left(\dfrac{4\pi R_i}{\lambda_i}\right)} = Q_{sys} + Q_{atm}$$

Where Q_{lpi} is the overall LPI quality factor in decibels, Q_{sys} is the clear weather system quality factor which is independent of water vapor, gas absorption, and rain attenuation, and the Q_{atm} element accounts for the relevant atmospheric components, as noted below.

$$Q_{atm} = LOSS_{Ri} - LOSS_{Rc}$$

$LOSS_{Ri}$ and $LOSS_{Rc}$ in the above equation are the intercept atmospheric path loss and the communications atmospheric path loss in decibels, respectively, and are functions of distance between the relative transmitter and receiver.

The goal is to integrate LPISDA (which calculates the Q_{sys}), and a weather model (which calculates Q_{atm}). An iterative approach is used to more accurately determine the maximum intercept receiver range R_i, and Q_{lpi}, given a maximum communication range R_c.

$$L_i = \left(20\log\left(\frac{4\pi R_c}{\lambda}\right) + LOSS_{Rc}\right) - Q_{sys} = \left(20\log\left(\frac{4\pi R_i}{\lambda_i}\right) + LOSS_{Ri}\right)$$

Equation 1

$$L_{i'} = 20\log\left(\frac{4\pi R_{i'}}{\lambda}\right) + LOSS_{Ri'}$$

Equation 2

1. Generate Q_{sys} from LPISDA.

2. Select maximum value of R_c, compute $LOSS_{Rc}$, and compute L_i from Equation 1.

3. Initialize $R_{i'}$ and compute $LOSS_{Ri'}$.

4. Solve for Li' in Equation 2, and compute $|L_i\text{-}L_i'| = \varepsilon$.

5. Repeat steps 3-4, perturbing $R_{i'}$ until $\varepsilon \le$ some ε_{max}.

6. Compute $Q_{lpi} = 20\log(R_{cmax}/R_{imax})$.

This algorithm can be used to calculate[3] the maximum intercept receiver range. It takes into account the atmospheric disturbances and presents an iterative solution to the problem.

<u>Remarks</u>

The difficulty of this interpretation is to find a weather transmissivity simulation which can be readily interfaced with LPISDA. LPISDA has subroutines which allow for this type of expansion. Several different models were approached to fill this task and most were found to be inadequate in one or more ways. Below is a list of the systems, and the specific attributes of each.

- *LOWTRAN:* The <u>Low</u> Resolution <u>Tran</u>smittance Code was developed in the early 1970's. LOWTRAN originated the AF family of codes, and it is essential to acknowledge it's

limitations and adopt more current and accurate algorithms [2]. LOWTRAN has slow execution time, is not coded with current coding standards, has low resolution, and was primarily designed for use in the optical frequency band. LOWTRAN cannot be used for frequencies below 30 GHz. For these reasons, it was not considered.

- *MODTRAN:* Moderate Resolution Transmittance Code, derived from LOWTRAN in 1989. MODTRAN was coded more efficiently and had higher resolution than LOWTRAN. Much like LOWTRAN, MODTRAN cannot be used for frequencies below 30 GHz. Coded in FORTRAN, manipulation of text input cards was excessively inefficient, and FORTRAN coding made integration very clumsy. It was agreed that MODTRAN was not an ideal integration component for LPISDA.

- *FASCODE:* Fast Atmospheric Signature Code, developed in 1978. Coded in FORTRAN, FASCODE is not particularly user friendly and is computer time and storage intensive. For this reason, integration of FASCODE into LPISDA would not be suitable for a PC implementation. FASCODE is usually maintained off-line for validation of more pragmatic codes [3].

- CRANE [4] : Named after Robert K. Crane of University of Oklahoma, this model can be used to calculate the atmospheric attenuation at frequencies between 1-100 GHz. This is the recommended model for adding atmospheric attenuation to the LPISDA. However, there are some problems with it, such as statistically important factors[4] which might be weakly represented in Crane's model. But, these various statistical errors are approximately of the same order of magnitude as statistical errors in LPISDA input parameters, so the CRANE model is deemed acceptable.

The CRANE model was determined to be the model of choice. The CRANE model uses rain belts which are empirically derived from many years of atmospheric precipitation recording. A table of values are given for each belt which gives the percentage of time that the

---

[3]Within a specific error tolerance.

rain fall will equal or exceed the expected rain rate. These rain rate tables, along with the communication and intercept receiver ranges and altitudes, are fed into the CRANE model equations, which derive the atmospheric attenuation of the communication paths and intercept receivers. A set of algorithms for computer implementation of a modified version of the CRANE model have been developed by the Air Force Information Warfare Center, Kelly AFB, Texas.

Before the location of the CRANE model, the author attempted to integrate MODTRAN into the LPISDA model. Because of the format of the FORTRAN records used as input to the MODTRAN model, the integration was inefficient. A search for a better model was instigated, and the CRANE model was found.

Conclusion

To make the simulation as user friendly as possible, the integration of the atmospheric transmissivity simulation code should be as transparent as possible. Ideally, the routines could be implemented as C subroutines which would allow for a seamless integration. However, most of the currently available products are written in FORTRAN, which is not compatible with the current LPISDA C language source code. If implementation in C cannot be accomplished, then the atmospheric transmissivity simulation source code should be compiled into executable code which can be called by the main LPISDA program. This solution would unfortunately be more tedious because it is easier to integrate source code written in the same language than it is to integrate separate pieces of executable code.

Throughout this project, it has been repeated that the best method of integration of LPISDA and a weather model would be to blend these systems into one format. The problem with this approach is the atmospheric simulation code must be both rewritten in C and verified. This integrated approach is still recommended because the final product which would

---

[4]The statistical representation of cloud layers, dynamics of various cloud systems, rain duration and such things.

result would be more effective than a system which was comprised of several independently written pre-compiled modules.

References

[1] *Low Probability of Intercept Communications System Performance Evaluation*, Lawrence L. Gutman, Fourth Government Organized LPI Exchange (GOLPIE) Conference, 2-4 June, 1993.

[2] *History of One Family of Atmospheric Radiative Transfer Codes*, G. P. Anderson, The European Symposium on Satellite Remote Sensing, Conference on Passive IP Remote Sensing of Clouds and Atmosphere, 26-30 Sep 1994.

[3] *Important New Features in PL/Geophysics Directorate Atmospheric Propagation Codes*, G.P. Anderson et. al., Joint Wright/Phillips Laboratories Atmospheric Workshop, 14 Nov 1995.

[4] *Prediction of Attenuation By Rain*, Robert K. Crane, IEEE Transactions On Communications Vol. Com-28, No. 9, September 1980.

Error Propagation in Decomposition
of Mueller Matrices

Diana M. Hayes
Graduate Student Teaching Fellow
Department of Mathematics

University of North Texas
Denton, TX 76203

Error Propagation in Decomposition of Mueller Matrices

Diana M. Hayes
Graduate Student Teaching Fellow
Department of Mathematics
University of North Texas

Abstract

A decomposition for Mueller matrices into three physically descriptive components was recently developed by Shih-Yau Lu [Lu, 1995, Lu and Chipman, 1996]. The effect of experimental error on this decomposition was studied. Both analytical and numerical methods were employed. Symbolic expression of the component matrices in terms of the original Mueller matrix elements shows how errors in the original matrix propogate through the decomposition. Complete symbolic decomposition was given for nondepolarizing Mueller matrices and their associated physical parameters; however, the depolarizing case produced unmanageably large expressions, so approximations were used. For the numerical results, Mathcad™ was used to randomly generate Mueller matrices, incorporate errors, and to analyze the effect of these errors on the decomposition. Results indicate that the error within the component matrices is proportional to the original error within the measured Mueller matrix, and that the proportional constant increases with each subsequent step in the decomposition. In addition, Cloude's method for eliminating "noise" in a Mueller matrix was employed, and its effect on error distribution was analyzed [Cloude, 1986, Cloude, 1989].

Error Propagation in Decomposition
of Mueller Matrices

Diana M. Hayes

1. Introduction

The Mueller-Stokes formalism is a method for describing the polarization properties of light and of optical systems. Polarization characteristics of light can be described by a four element vector, called a Stokes vector. Stokes vectors have the form $\begin{pmatrix} s_0 & s_1 & s_2 & s_3 \end{pmatrix}$ where s_0 describes the overall intensity of the light, and s_1, s_2, and s_3 describe horizontal, 45 degree, and circular polarization, respectively. An optical system which translates one state of polarized light into another can be represented by a 4x4, real valued Mueller matrix. Since errors can occur when gathering physical data, the resultant matrix may not represent the actual Mueller matrix for an optical system. The influence of these errors on the scientist's analysis of the optical system is the main focus of this paper.

In particular, this paper describes the effect of experimental error on the decomposition of Mueller matrices described by Shih-Yau Lu [Lu, 1995]. In his dissertation Lu developed the decomposition of a Mueller matrix into the product of a depolarizer, a retarder, and a diattenuator. If a Mueller matrix is computed experimentally then the accuracy of the elements within the original matrix determines the accuracy of each component matrix, and ultimately affects the scientist's analysis of the optical system. Determining how error within the original matrix is propagated through Lu's decomposition process is of use to the researcher for interpreting the validity of his results and is the primary topic of this report.

2. Decomposition of Mueller Matrices [Lu 1995, Lu & Chipman 1995]

Mueller matrices are divided into two categories, nondepolarizing and depolarizing matrices. As their name indicates, nondepolarizing Mueller matrices are associated with optical systems that do not scatter incident polarized light. These nondepolarizing Mueller matrices can be written as the product of a retarder and a diattenuator, $M = M_R M_D$. Lu adopts a 2x2 shorthand notation for these matrices which proves useful for their decomposition. Let M be the Mueller matrix scaled so that the upper left corner is one.

Then M has the form

$$M = \begin{bmatrix} 1 & m_{01} & m_{02} & m_{03} \\ m_{10} & m_{11} & m_{12} & m_{13} \\ m_{20} & m_{21} & m_{22} & m_{23} \\ m_{30} & m_{31} & m_{32} & m_{33} \end{bmatrix} = \begin{bmatrix} 1 & \overline{D}^{\mathrm{T}} \\ \overline{P} & m \end{bmatrix} \quad \text{with the lower 3x3,} \quad m = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix}.$$

The first row, $\overline{D}$, denotes the diattenuation vector and the first column, $\overline{P}$, denotes the polarizance vector,

$$\overline{D} = \begin{bmatrix} D_{\mathrm{H}} \\ D_{45} \\ D_{\mathrm{C}} \end{bmatrix} = \begin{bmatrix} m_{01} \\ m_{02} \\ m_{03} \end{bmatrix} \quad \text{and} \quad \overline{P} = \begin{bmatrix} P_{\mathrm{H}} \\ P_{45} \\ P_{\mathrm{C}} \end{bmatrix} = \begin{bmatrix} m_{10} \\ m_{20} \\ m_{30} \end{bmatrix}.$$

The subscripts represent the horizontal, 45 degree, and circular components of each vector.

To accomplish the decomposition, the diattenuator, M_{D}, is calculated from the first row of M, then M_{D}^{-1} is multiplied to M to yield the retarder, $M_{\mathrm{R}} = M M_{\mathrm{D}}^{-1}$. The form of the diattenuator matrix is given by

$$M_{\mathrm{D}} = \begin{bmatrix} 1 & \overline{D}^{\mathrm{T}} \\ \overline{D} & m_{\mathrm{D}} \end{bmatrix} \quad \text{where the lower 3x3} \quad m_{\mathrm{D}} = a\mathbf{I}_3 + b(\overline{D} \cdot \overline{D}^{\mathrm{T}}),$$

and where a and b are scalars derived from the norm of the diattenuation vector,

$$D = |\overline{D}|, \qquad a = \sqrt{1 - D^2}, \quad \text{and} \quad b = \frac{1 - \sqrt{1 - D^2}}{D^2}.$$

Inverting M_{D} can be accomplished numerically or by using the expression

$$M_{\mathrm{D}}^{-1} = \frac{1}{a^2} \begin{bmatrix} 1 & -\overline{D}^{\mathrm{T}} \\ -\overline{D} & \mathbf{I}_3 \end{bmatrix} + \frac{1}{a^2(a+1)} \begin{bmatrix} 0 & \overline{0}^{\mathrm{T}} \\ \overline{0} & (\overline{D} \cdot \overline{D}^{\mathrm{T}}) \end{bmatrix}.$$

The remaining retarder component has the form,

$$M_{\mathrm{R}} = \begin{bmatrix} 1 & \overline{0}^{\mathrm{T}} \\ 0 & m_{\mathrm{R}} \end{bmatrix} \quad \text{where the lower 3x3 } m_{\mathrm{R}} \text{ is a rotation matrix.}$$

For nondepolarizing Mueller matrices, this lower 3x3 is given by $m_{\mathrm{R}} = \frac{1}{a}\left[m - b(\overline{P} \cdot \overline{D}^{\mathrm{T}})\right].$

The retardance, R, and the retardance vector (the fast axis for the retarder component), $\overline{R}$, can be calculated from the retarder matrix by the equations

$$R = |\overline{R}| = \cos^{-1}\left(\frac{tr(m_R) - 1}{2}\right) \quad \text{and} \quad \overline{R} = \begin{bmatrix} R_H \\ R_{45} \\ R_C \end{bmatrix} = \begin{bmatrix} (M_R)_{23} - (M_R)_{32} \\ (M_R)_{31} - (M_R)_{13} \\ (M_R)_{12} - (M_R)_{21} \end{bmatrix} \frac{R}{2\sin(R)} .$$

These formulas for retardance and the retardance vector apply to either nondepolarizing or depolarizing Mueller matrices. If retardance is expected to be in the interval $(\pi, 2\pi)$ then retardance can be calculated by

$$R = 2\pi - \cos^{-1}\left(\frac{tr(m_R) - 1}{2}\right).$$

Depolarizing Mueller matrices can be decomposed into a product of three components, a depolarizer, a retarder, and a diattenuator, $M = M_\Delta M_R M_D$. The first step in the decomposition is the same as in the nondepolarizing case; compute the diattenuator and eliminate it, leaving the composition $M' = MM_D^{-1} = M_\Delta M_R$. The 2x2 notation is given by

$$M_\Delta M_R = \begin{bmatrix} 1 & \overline{0}^T \\ \overline{P}_\Delta & m_\Delta \end{bmatrix}\begin{bmatrix} 1 & \overline{0}^T \\ \overline{0} & m_R \end{bmatrix} = \begin{bmatrix} 1 & \overline{0}^T \\ \overline{P}_\Delta & m_\Delta m_R \end{bmatrix} = \begin{bmatrix} 1 & \overline{0}^T \\ \overline{P}_\Delta & m' \end{bmatrix} = M' .$$

Noting that the transpose of the retarder is also its inverse and that the lower 3x3 of the depolarizer, m_Δ, is symmetric yields an expression for m_Δ^2. The expressions and computations using the lower 3x3 of the corresponding matrices are

$$m_R = m_R^{-1} , \text{ and } m_\Delta = m_\Delta^T \quad \text{imply} \quad m'(m')^T = m_\Delta m_R (m_\Delta m_R)^T = m_\Delta^2 .$$

From this result the eigenvalues of m_Δ can be calculated, which yield an expression for m_Δ using the Cayley-Hamilton theorem. This part of the decomposition process is discussed in detail during later sections of this report. To complete the construction of the depolarizer, the polarizance vector of M_Δ is computed by the equation, $\overline{P}_\Delta = \frac{1}{a^2}(\overline{P} - m\overline{D})$. M_Δ is then inverted and applied to M' to yield the retarder, $M_R = M_\Delta^{-1} M'$. Retardance can be calculated by the same methods as in the nondepolarizing case.

As previously mentioned, this decomposition and the notation involved were developed by Lu and Chipman [Lu, 1995; Lu and Chipman, 1995]. A more thorough discussion is given in these references. The next two sections of this paper contain the symbolic decomposition of Mueller matrices in terms of the original matrix elements. Writing the component matrices in

this form displays how error within the original Mueller matrix propagates throughout the decomposition.

3. Decomposition of Nondepolarizing Mueller Matrices

Recall that nondepolarizing Mueller matrices have the decomposition $M = M_R M_D$, and that the original Mueller matrix was denoted by

$$M = \begin{bmatrix} 1 & m_{01} & m_{02} & m_{03} \\ m_{10} & m_{11} & m_{12} & m_{13} \\ m_{20} & m_{21} & m_{22} & m_{23} \\ m_{30} & m_{31} & m_{32} & m_{33} \end{bmatrix}.$$

The diattenuator component has the form

$$M_D = \begin{bmatrix} 1 & m_{01} & m_{02} & m_{03} \\ m_{01} & a + b(m_{01})^2 & b(m_{01}m_{02}) & b(m_{01}m_{03}) \\ m_{02} & b(m_{02}m_{01}) & a + b(m_{02})^2 & b(m_{02}m_{03}) \\ m_{03} & b(m_{03}m_{01}) & b(m_{03}m_{02}) & a + b(m_{03})^2 \end{bmatrix},$$

where $a = \sqrt{1 - (m_{01}{}^2 + m_{02}{}^2 + m_{03}{}^2)}$ and $b = \dfrac{1 - \sqrt{1 - (m_{01}{}^2 + m_{02}{}^2 + m_{03}{}^2)}}{(m_{01}{}^2 + m_{02}{}^2 + m_{03}{}^2)}$.

Furthermore, the diattenuation vector is given simply by the first row of M,

$$\overline{D} = \begin{bmatrix} D_H \\ D_{45} \\ D_C \end{bmatrix} = \begin{bmatrix} m_{01} \\ m_{02} \\ m_{03} \end{bmatrix}, \text{ and total diattenuation } D = \sqrt{m_{01}{}^2 + m_{02}{}^2 + m_{03}{}^2}.$$

Hence, error within the diattenuator is solely determined by errors within the first row of M. The retarder matrix in the nondepolarizing case is expressed as

$$M_R = \frac{1}{a} \begin{bmatrix} a & 0 & 0 & 0 \\ 0 & m_{11} - b(m_{10}m_{01}) & m_{12} - b(m_{10}m_{02}) & m_{13} - b(m_{10}m_{03}) \\ 0 & m_{21} - b(m_{20}m_{01}) & m_{22} - b(m_{20}m_{02}) & m_{23} - b(m_{20}m_{03}) \\ 0 & m_{31} - b(m_{30}m_{01}) & m_{32} - b(m_{30}m_{02}) & m_{33} - b(m_{30}m_{03}) \end{bmatrix}.$$

Hence, in the nondepolarizing case, the overall retardance is given by

$$R = \cos^{-1}\left(\frac{1}{2a} \big[m_{11} + m_{22} + m_{33} - b(m_{10}m_{01} + m_{20}m_{02} + m_{30}m_{03}) - a \big] \right),$$

and the retardance vector has the form

$$\overline{R} = \begin{bmatrix} m_{23} - m_{32} - b(m_{20}m_{03} - m_{30}m_{02}) \\ m_{31} - m_{13} - b(m_{30}m_{01} - m_{10}m_{03}) \\ m_{12} - m_{21} - b(m_{10}m_{02} - m_{20}m_{01}) \end{bmatrix} \frac{\cos^{-1}\left(\frac{1}{2a}\left[m_{11} + m_{22} + m_{33} - b(m_{10}m_{01} + m_{20}m_{02} + m_{30}m_{03}) - a\right]\right)}{\sqrt{4a^2 - \left[m_{11} + m_{22} + m_{33} - b(m_{10}m_{01} + m_{20}m_{02} + m_{30}m_{03}) - a\right]^2}}$$

4. Decomposition of Depolarizing Mueller Matrices

The decomposition process for depolarizing Mueller matrices begins with the same construction of the diattenuator matrix as related in the previous case. Hence, the errors within the diattenuator component and the diattenuation vector are completely determined by errors in the first row of M. Eliminating M_D leaves $M' = MM_D^{-1} = M_\Delta M_R$. Calculating the depolarizer, M_Δ, is the next step. The lower 3x3, m_Δ, is symmetric, so its eigenvalues are real. Calculating the eigenvalues of $m'(m')^T = m_\Delta m_R (m_\Delta m_R)^T = m_\Delta^2$ yields the squares of the eigenvalues of m_Δ. However, there is some ambiguity for the signs of the eigenvalues of m_Δ. The retarder is a rotation matrix and has positive determinant, so the sign of the determinant of m' indicates the sign of the determinant of m_Δ. If m' has positive determinant, then all three eigenvalues could be positive, or any two could be negative. Similarly, if m' has negative determinant then all three eigenvalues could be negative or any one of the three could be negative. Lu assumes that all three eigenvalues have the same sign, which simplifies the expression for m_Δ, and allows only one possibility depending on the sign of the determinant of m'. It should be noted that the choice of signs for the eigenvalues does affect the remaining retarder matrix, and in general will change the resulting retardance and the retardance vector. However, for most laboratory experiments where depolarization effects of the optical system are small, then these eigenvalues are all close to one and Lu's assumption is valid. Lu's expression for m_Δ is derived from the Cayley-Hamilton theorem, which states that a matrix is a root of its characteristic polynomial. If p^2, q^2, and r^2 are the eigenvalues for m_Δ^2, where p, q, and r are positive, then p, q, and r denote the absolute values of the eigenvalues for m_Δ.

Let $\qquad K_1 = p + q + r \qquad K_2 = pq + pr + qr \qquad K_3 = pqr$.

Applying the Cayley-Hamilton theorem yields

$$m_\Delta = \pm \left[m'(m')^T + K_2 I \right]^{-1} \left[K_1 m'(m')^T + K_3 I \right],$$

where the sign in front is equal to the sign of the determinant of m'. Applying m_Δ^{-1} to m' gives a symbolic representation for the retarder

$$m_\Delta^{-1} m' = m_R = \pm \left[K_1 m'(m')^T + K_3 I \right]^{-1} \left[m'(m')^T m' + K_2 m' \right].$$

Theoretically, p, q, and r can be calculated in terms of the original Mueller matrix elements by using the algebraic solution to a cubic polynomial. Unfortunately, these expressions are so long and complicated, that they are of little practical use. In light of this fact, determining or at least approximating the K's will prove more feasible. K_3 can be calculated directly using the following result.[For related results see Gil and Bernabeu 1985, Kostinski 1992, and Simon 1981.]

$$K_3 = \det(m_\Delta) = \sqrt{\det(m_\Delta^2)} = \sqrt{\det(m'(m')^T)} = \det(m').$$

Recall $M' = M(M_\Delta)^{-1}$ has the form $M' = \begin{bmatrix} 1 & \overline{0}^T \\ \overline{P_\Delta} & m' \end{bmatrix}$, and so

$$K_3 = \det(m') = \det(M') = \det(M) \det(M_\Delta^{-1}) = \frac{\det(M)}{\det(M_\Delta)} = \frac{\det(M)}{a^4}.$$

Let $T = trace[m_\Delta^2] = p^2 + q^2 + r^2$, and $t = trace\left[K_3^2 (m_\Delta^2)^{-1} \right] = p^2 q^2 + p^2 r^2 + q^2 r^2$. Then

K_1 satisfies the recursive equation $K_1 = \sqrt{T + 2\sqrt{t + 2 K_3 K_1}}$, and can be approximated by

$K_1 \approx \sqrt{T + 2\sqrt{t + 2 K_3 \sqrt{T}}}$. Noting that $K_2 = \frac{1}{2}\left[K_1^2 - T \right]$ gives the approximation

$K_2 \approx \sqrt{t + 2 K_3 \sqrt{T}}$. Below is the expression for T in terms of the original Mueller matrix elements, and the expression for t in terms of m_Δ^2 as defined below.

$$T = \frac{1}{a^2} \left[\sum_{i,j=1}^{3} m_{i,j}^2 - \sum_{i=1}^{3} m_{i,0}^2 \right] + \frac{1}{a^4} \left[\sum_{i=1}^{3} \left(m_{i,0} - \sum_{j=1}^{3} m_{i,j} m_{0,j} \right)^2 \right], \text{ and}$$

$$t = \Delta_{2,2} \Delta_{3,3} + \Delta_{1,1} \Delta_{3,3} + \Delta_{1,1} \Delta_{2,2} - \left(\Delta_{2,3}^2 + \Delta_{1,3}^2 + \Delta_{1,2}^2 \right).$$

The retarder matrix is the last component to be uncovered in this decomposition process. Since $m_\Delta^{-1}m' = m_R$, the following expression for m_Δ^{-1} as a function of m_Δ^2 may be useful, especially if the depolarizer eigenvalues p, q, and r can be approximated.

$$m_\Delta^{-1} = \frac{1}{K_1}\left[\alpha\mathbf{I} - \beta m_\Delta^2 + \gamma K_3^2\left(m_\Delta^2\right)^{-1}\right] \quad \text{where} \quad \beta = \frac{1}{(p+q)(p+r)(q+r)} ,$$

$$\gamma = \frac{(p+q+r)}{(pqr)(p+q)(p+r)(q+r)} , \quad \text{and} \quad \alpha = \frac{\left(p^2+q^2+r^2\right)(p+q+r) - pqr}{(p+q)(p+r)(q+r)} + 1 .$$

If m_Δ^2 is given the notation $m_\Delta^2 = \begin{bmatrix} \Delta_{1,1} & \Delta_{1,2} & \Delta_{1,3} \\ \Delta_{2,1} & \Delta_{2,2} & \Delta_{2,3} \\ \Delta_{3,1} & \Delta_{3,2} & \Delta_{3,3} \end{bmatrix}$ (note that $\Delta_{i,j} = \Delta_{j,i}$) where

$$\Delta_{i,j} = \frac{1}{a^2}\left[\left(\sum_{k=1}^{3} m_{ik}m_{jk}\right) - m_{i0}m_{j0}\right] + \frac{1}{a^4}\left[m_{i0} - \sum_{k=1}^{3} m_{ik}m_{0k}\right]\left[m_{j0} - \sum_{k=1}^{3} m_{jk}m_{0k}\right], \quad \text{then matrix}$$

$$K_3^2\left(m_\Delta^2\right)^{-1} = \begin{bmatrix} \Delta_{2,2}\Delta_{3,3} - \Delta_{2,3}^2 & \Delta_{1,3}\Delta_{2,3} - \Delta_{1,2}\Delta_{3,3} & \Delta_{1,2}\Delta_{2,3} - \Delta_{2,2}\Delta_{2,3} \\ \Delta_{1,3}\Delta_{2,3} - \Delta_{1,2}\Delta_{3,3} & \Delta_{1,1}\Delta_{3,3} - \Delta_{1,3}^2 & \Delta_{1,2}\Delta_{1,3} - \Delta_{1,1}\Delta_{2,3} \\ \Delta_{1,2}\Delta_{2,3} - \Delta_{2,2}\Delta_{2,3} & \Delta_{1,2}\Delta_{1,3} - \Delta_{1,1}\Delta_{2,3} & \Delta_{1,1}\Delta_{2,2} - \Delta_{1,2}^2 \end{bmatrix} .$$

So the retarder rotation matrix is given by $m_R = m_\Delta^{-1}m' = \frac{1}{K_1}\left[\alpha\mathbf{I} - \beta m_\Delta^2 + \gamma K_3^2\left(m_\Delta^2\right)^{-1}\right]m'$.

5. Error Distribution in Physical Parameters

The previous sections have dealt with the symbolic representation of the physical parameters of diattenuation, retardance, and depolarization from Lu's decomposition. These expressions show which terms in the Mueller matrix are involved in calculating each component. If errors are present in the original Mueller matrix, then they filter through the decomposition and show up as error within these physical parameters. Since the expressions for parameters become increasingly complicated as the decomposition progresses, the net error is not easily determined. To get an indication of how the distribution of error in the original matrix affects the distribution of error within diattenuation, retardance, and depolarization, a large sample of numerical examples were constructed and analyzed. Mathcad™ was used to randomly generate Mueller matrices, introduce error, and to analyze the results. The Mueller matrices were generated using the reverse of Lu's decomposition. A randomly generated matrix of errors was then added to each Mueller matrix. The error within each element of the matrix was normally distributed with mean $\mu = 0$ and standard deviation $\sigma = .003$. This choice was determined by taking multiple runs from a matrix polarimeter and calculating the standard deviation within each element.

It is important to note that Lu's decomposition process applies only to physically realizable Mueller matrices, and after these random errors were added, the matrices in our experiment no longer satisfied this condition. Hence, Cloude's method for finding the closest physical Mueller matrix to a given 4x4 matrix was employed [Cloude, 1989]. The following statistical results include an analysis of the effect of Cloude's method on the original error distribution.
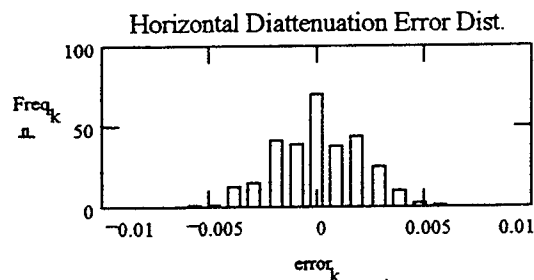
Sample 1) Nondepolarizing Mueller Matrices

Recall that a nondepolarizing Mueller matrix is decomposed into the product of a retarder and a diattenuator, and thus, the information regarding retardance and diattenuation is restricted to the corresponding matrix. The diattenuation vector is read directly from the first row of the original Mueller matrix, so tracing errors in diattenuation is straightforward. Hence, the resulting error distribution for the diattenuation vector components shows the effect of Cloude's method for "noise" reduction. Recall the original error introduced was normally distributed with mean $\mu = 0$ and standard deviation $\sigma = .003$. The results from this experiment show a decrease in standard deviation in these first row entries to approximately $\sigma = .002$. This decrease implies that Cloude's method is effective at reducing noise in an experimental nondepolarizing Mueller matrix.

A sample of 300 Mueller matrices were generated to give the error distributions shown on the following page. The first graph shows the error in overall diattenuation D, which is the norm of the diattenuation vector. The error distribution for each of the three components in the vector follows. The values for the mean and standard deviation of each distribution are also listed. Retardance and the three retardance vector components were handled in a similar fashion. However, the standard deviation listed for the three retardance vector components, denoted by "s" instead of "σ", was calculated from the distribution shown. Each of these three distributions include approximately 93% of the data values. The remaining values, though a small percentage, were spread far enough to significantly bias the standard deviation. This indicates that relatively small errors in the original Mueller matrix can result in large errors within the retardance component approximately seven percent of the time. Multiple runs were performed using Mathcad™ to confirm this result.

Nondepolarizing Matrices -- Error Distributions
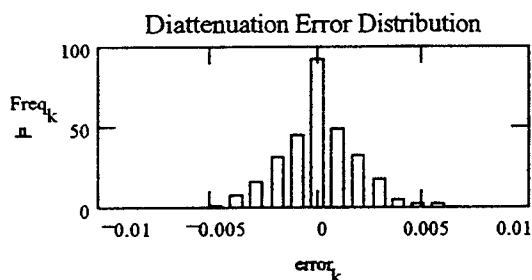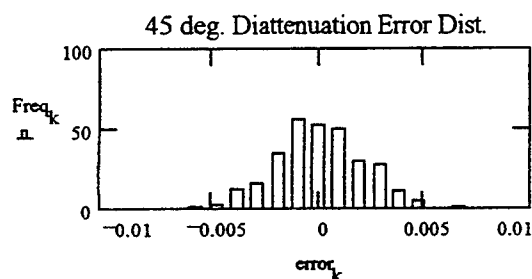
Diattenuation Error Distribution

$\mu_D = 0$ $\qquad$ $\sigma_D = 0.0017$

Horizontal Diattenuation Error Dist.

$\mu_{Dh} = 0.0001$ $\qquad$ $\sigma_{Dh} = 0.002$

45 deg. Diattenuation Error Dist.

$\mu_{D45} = 0.0001$ $\qquad$ $\sigma_{D45} = 0.0021$

Circular Diattenuation Error Dist.

$\mu_{Dc} = 0$ $\qquad$ $\sigma_{Dc} = 0.0018$

Retardance Error Distribution

$\mu_R = 0$ $\qquad$ $\sigma_R = 0.0028$

Horizontal Retardance Error Distribution

$\mu_{Rh} = -0.0004$ $\qquad$ $s_{Rh} = 0.0053$

45 degree Retardance Error Distribution

$\mu_{R45} = -0.0006$ $\qquad$ $s_{R45} = 0.0051$

Circular Retardance Error Distribution
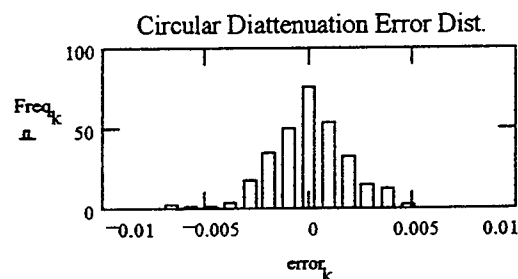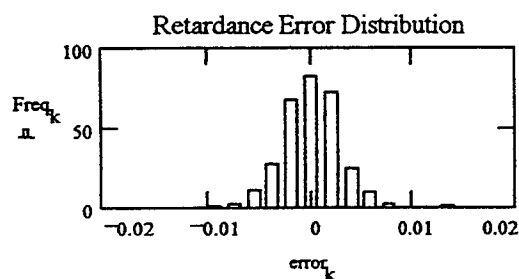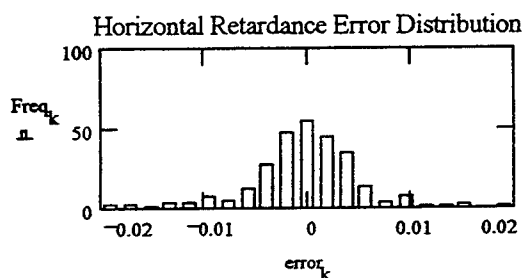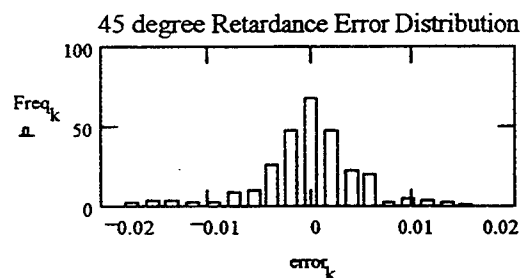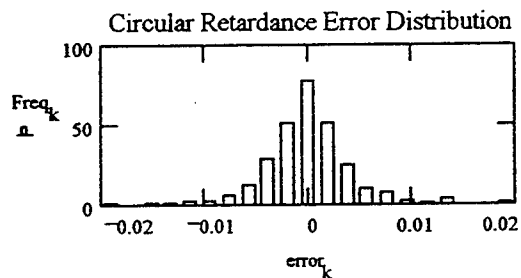
$\mu_{Rc} = 0.0003$ $\qquad$ $s_{Rc} = 0.0043$

Sample 2) Depolarizing Mueller Matrices

A depolarizing Mueller matrix is the product of a retarder, depolarizer, and diattenuator, so there are three basic parameters to be measured in this decomposition. Retardance and diattenuation are measured as before, and the depolarization is measured by the quantity Δ. Cloude's method was again employed, but it was found to be ineffective in reducing noise in the depolarizing case. Also note that since the retarder is the last matrix to be derived from Lu's decomposition, the standard deviation of the error distribution was larger than in the nondepolarizing case.

A sample of 300 matrices were generated and analyzed. The error distributions are given below. During the course of this computer simulation it was discovered that when depolarization was large, the errors in depolarization and in retardance were often very large (i.e. close to the maximum possible error). The reason for these large errors was directly related to the change in sign of the eigenvalues of the depolarizer matrix. When these eigenvalues are close to zero, relatively small errors in the original matrix can change the sign of these eigenvalues. For instance, moving the eigenvalue 0.1 to -.05 is only a change of -0.15 units, but the change in sign will drastically affect the retarder matrix given as the last step in Lu's decomposition. For this reason, the depolarizer component matrices in the following sample were restricted to having eigenvalues of absolute value greater than 0.5 . A description of the structure of a depolarizer Mueller matrix in terms of its eigenvalues and polarizance vector is included in the appendix. The mean and standard deviation of the entire sample are shown for diattenuation, and the sample standard deviation derived from the distribution shown is given for retardance and depolarization. "T" denotes the number of matrices included in the truncated distributions.

Depolarizing Matrices — Error Distributions
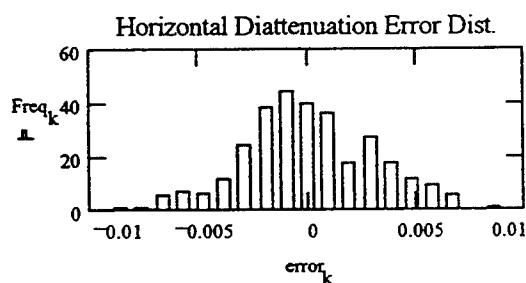
Diattenuation Error Distribution

Horizontal Diattenuation Error Dist.

$\mu_D = -0.0006$ $\sigma_D = 0.0033$ $\mu_{Dh} = 5.6964 \cdot 10^{-5}$ $\sigma_{Dh} = 0.0029$

45 deg. Diattenuation Error Dist.

$\mu_{D45} = -2.0824 \cdot 10^{-6}$　　　$\sigma_{D45} = 0.0031$

Circular Diattenuation Error Dist.

$\mu_{Dc} = 2.8861 \cdot 10^{-5}$　　　$\sigma_{Dc} = 0.0032$

Retardance Error Distribution

$\mu_R = -0.0011$　$s_R = 0.0061$　$T_R = 296$

Horizontal Retardance Error Distribution
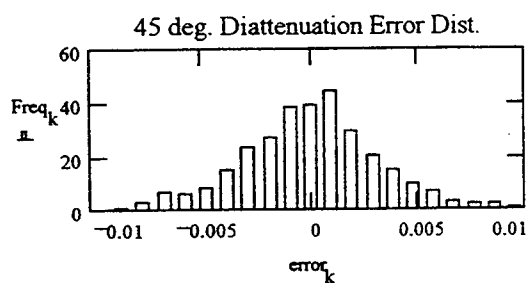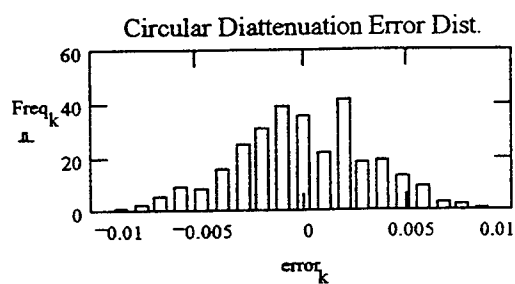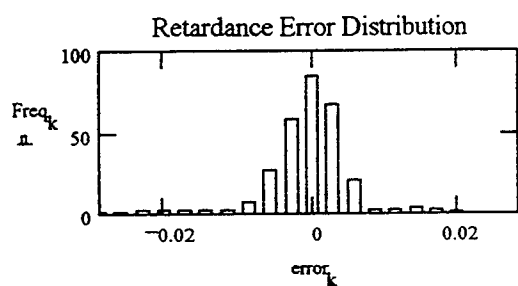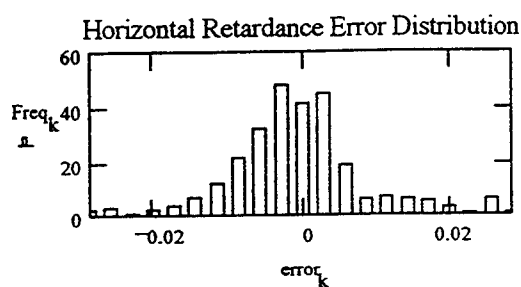
$\mu_{Rh} = -0.0093$　$s_{Rh} = 0.0126$　$T_{Rh} = 273$

45 degree Retardance Error Distribution
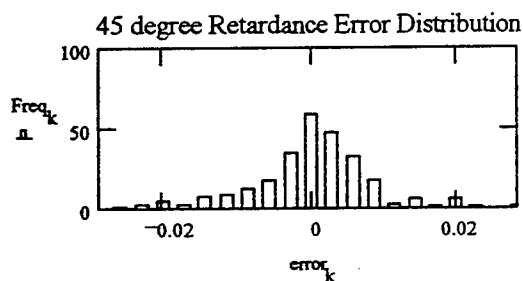
$\mu_{R45} = 0.0044$　$s_{R45} = 0.009$　$T_{R45} = 265$

Circular Retardance Error Distribution
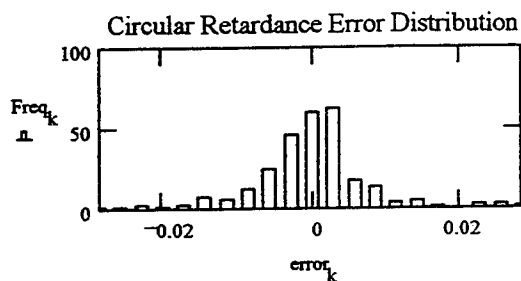
$\mu_{Rc} = 0.0172$　$s_{Rc} = 0.0189$　$T_{Rc} = 275$

Depolarization Error Dist.
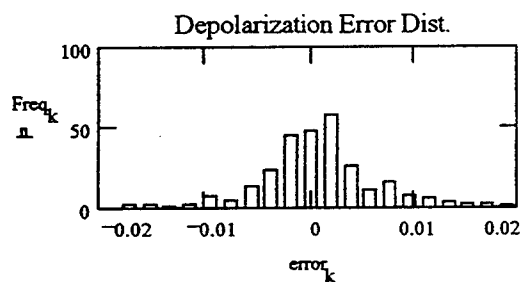
$\mu_\Delta = 0.0012$　$s_\Delta = 0.0057$　$T_\Delta = 285$

A second series of Mueller matrices were generated to determine the correlation between error in the parameters after decomposition and error within the original Mueller matrix. A series of non depolarizing Mueller matrices were generated. Twenty nine sets of 400 Mueller matrices each were generated; each set was then given normally distributed random errors. The standard deviations increased over the 29 sets, starting with $\sigma = .002$ and ending with $\sigma = .100$. Cloude's method of retaining only the largest eigenvalue from the coherency matrix to yield the closest nondepolarizing Mueller matrix was again employed. [Cloude 1989] The reduction in standard deviation within the diattenuation components implies that Cloude's method in the nondepolarizing case reduces error by approximately one third. Recall that the diattenuation components are read directly from the first row of the matrix, and show directly the effect of Cloude's method on the complete matrix. Only overall retardance was analyzed since it was clearly linearly related to error size, and since the three vector components were not so neatly behaved.

A similar series of depolarizing matrices were generated, and Cloude's method of retaining only the positive eigenvalues of the coherency matrix was employed. However, only a small reduction in error, approximately 10%, was seen within the diattenuation components. So Cloude's method has a slight effect on the original Mueller matrix. Note the increase in standard deviation with each step of the decomposition process. First diattenuation, then depolarization and finally retardance. It is important to note that the individual retardance components and the depolarization eigenvalues often varied widely from the original values. Therefore, overall retardance, R, and the parameter Δ, which "average" their respective individual components, were analyzed since they tend to smooth out the correlation. As in the previous analysis of depolarizing Mueller matrices, the size of the eigenvalues of the depolarizer components were again restricted between 0.5 and 1.

The graphs on the following pages show the standard deviation of the diattenuation, retardance, and depolarization parameters as functions of the standard deviation of the error introduced to the complete Mueller matrices. The 29 sample data points and the corresponding "least squares" line for each data set are pictured, and the equation of the line is written below the graph.

Nondepolarizing Mueller Matrices
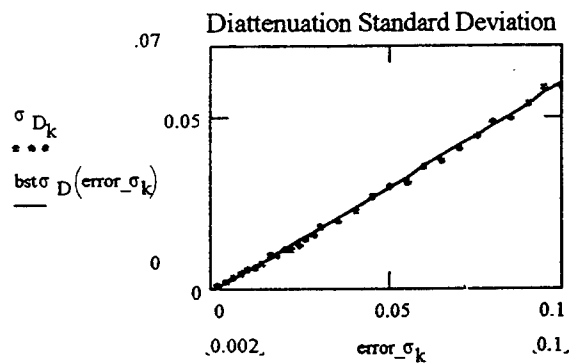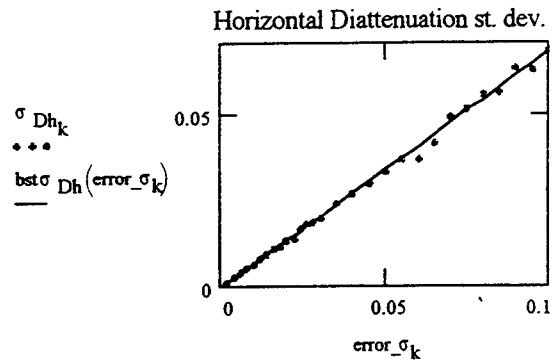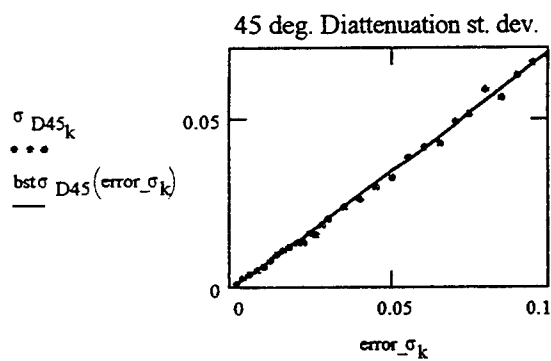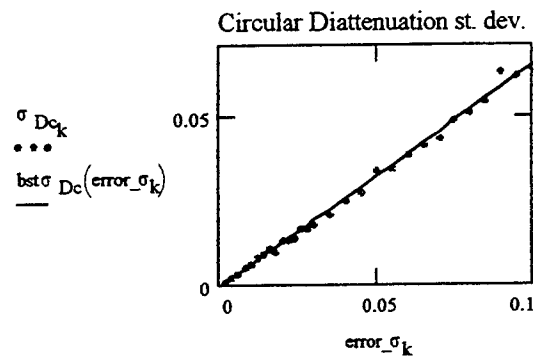
Diattenuation Standard Deviation

σ_{D_k}
• • •

$bst\sigma_D\left(error\_\sigma_k\right)$
___

$$bst\sigma_D(x) := .599x - .000\epsilon$$

Horizontal Diattenuation st. dev.

σ_{Dh_k}
• • •

$bst\sigma_{Dh}\left(error\_\sigma_k\right)$
___

$$bst\sigma_{Dh}(x) := .678\mathord{1}x - .000\mathord{?}$$

45 deg. Diattenuation st. dev.

σ_{D45_k}
• • •

$bst\sigma_{D45}\left(error\_\sigma_k\right)$
___

$$bst\sigma_{Dc}(x) := .6482x - .000\mathord{4}$$

Circular Diattenuation st. dev.

σ_{Dc_k}
• • •

$bst\sigma_{Dc}\left(error\_\sigma_k\right)$
___

$$bst\sigma_{D45}(x) := .6908x - .000\mathord{4}$$

Retardance standard deviation

$bstF\left(err\_\sigma_q\right)$
___

σs_{R_q}
• • •

$$bstF(x) := 1.0093x + .000\mathord{\mathsterling}$$

12-16

Depolarizing Mueller Matrices

Diattenuation standard deviation

σD_k
• • •
$line_k$ —

$line\sigma_D(x) := .9454x + .002\angle$

Horizontal Diattenuation st. dev.

σDh_k
• • •
$line_k$ —

$line\sigma_{Dh}(x) := .8666x + .002\angle$

45 degree Diattenuation st. dev.

$\sigma D45_k$
• • •
$line_k$ —

$line\sigma_{D45}(x) := .8787x + .002\angle$

Circular Diattenuation st. dev.

σDc_k
• • •
$line_k$ —

$line\sigma_{Dc}(x) := .9252x + .001\overline{}$

Depolarization standard deviation

0.13593
$s \Delta_k$
• • •
$line_k$ —
0.0039

0.002 0.1

$line\_s_\Delta(x) := 1.2573x + .010\overline{}$

Retardance standard deviation

$s R_k$
• • •
$line_k$ —
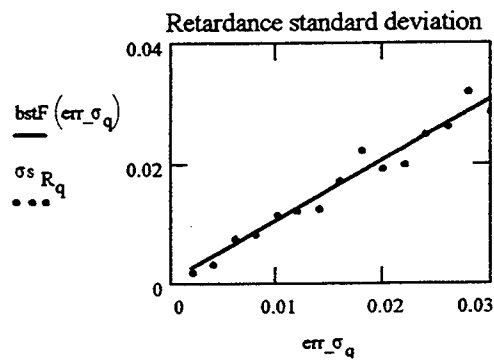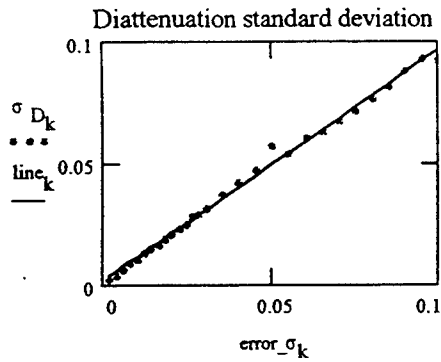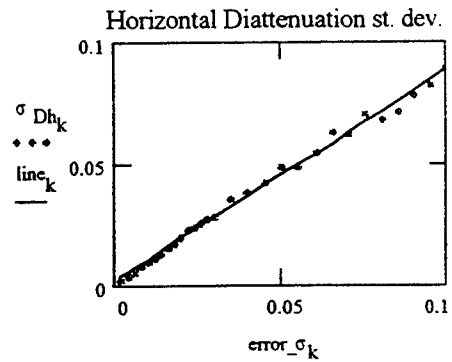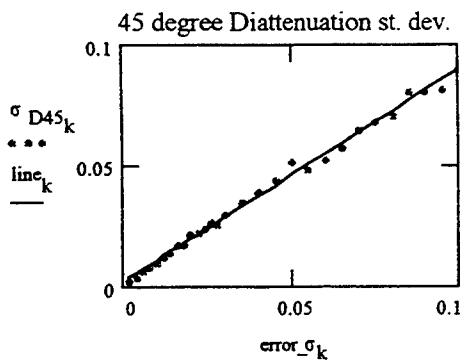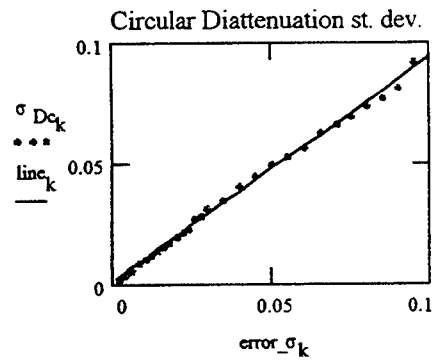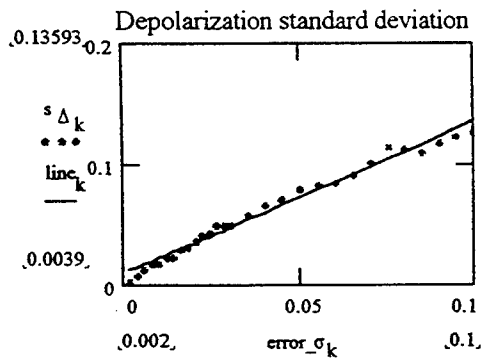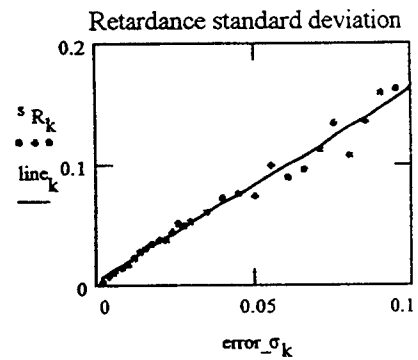
$line\_s_R(x) := 1.5852x + .004\overline{}$

6. Conclusions

As expected, error within an experimental Mueller matrix propagated through to the component matrices and the size of the error increased with each step in the decomposition. In particular, as the last step in the decomposition, the retarder matrix displayed the greatest sensitivity these experimental errors. However, even though the retardance vector elements varied widely as a result of this sensitivity, the overall retardance maintained a consistent linear correlation. The error within the diattenuation and depolarization parameters also clearly displayed a linear correlation to error within the original matrix. The depolarizing matrices with large depolarizing factors (eigenvalues close to zero) had to be eliminated in order to maintain the correlation in a large percentage of the sample (over 90%). Relatively small changes in the eigenvalues close to zero resulted in drastic alteration of the retardance parameters. Since the decomposition process was applicable only to "physical" Mueller matrices, Cloude's method was employed and seen to decrease the error size. Cloude's method was seen to be effective at reducing the noise in a nondepolarizing optical system by approximately one third, and approximately ten percent of the noise was reduced in a depolarizing system.

Appendix - Description of Pure Depolarizer Mueller Matrix

A pure depolarizer Mueller matrix was described by Lu as a matrix with the form

$$M_\Delta = \begin{bmatrix} 1 & \overline{0}^T \\ \overline{P}_\Delta & m_\Delta \end{bmatrix} \quad \text{where } \overline{P}_\Delta = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \text{ is the polarizance vector of the matrix, and where } m_\Delta \text{ is a}$$

3x3 real symmetric matrix with eigenvalues p, q, and r, each having absolute value less than or equal to one. This description is incomplete because such a matrix may not be classified as a physical Mueller matrix as defined by Cloude [1986, 1989]. More restrictions on the values for p, q, and r, and the polarizance vector components are needed. Cloude describes a physical Mueller matrix as a 4x4 real matrix whose corresponding coherency matrix has nonnegative eigenvalues.

$$\text{Taking the 4x4 matrix } M_\Delta = \begin{bmatrix} 1 & 0 & 0 & 0 \\ v_1 & p & 0 & 0 \\ v_2 & 0 & q & 0 \\ v_3 & 0 & 0 & r \end{bmatrix},$$

constructing its corresponding coherency matrix, and applying Fermat's "rule of signs" Theorem to the characteristic polynomial of the coherency matrix yields the following observations. The characteristic polynomial has leading coefficient 1 and degree three coefficient -4, so only three inequalities yield the restrictions on M_Δ.

$$\text{Let} \quad P = v_1^2 + v_2^2 + v_3^2 , \quad A = p^2 + q^2 + r^2 , \quad \text{and}$$

$$\mathcal{D} = (1 + p + q + r)(1 + p - q - r)(1 - p + q - r)(1 - p - q + r).$$

$$A + P \leq 3 ,$$

$$A + P \leq 1 + 2pqr$$

$$\mathcal{D} + P^2 + 2\left[AP - 2\left(p^2 v_1^2 + q^2 v_2^2 + r^2 v_3^2 \right) \right] \geq 0$$

Furthermore, van der Mee [1989] proved a theorem about the eigenvalues of a physical Mueller matrix which implies that each of the four factors of $\mathcal{D}$ must be nonnegative [see also Cloude, 1989, and Girgel', 1991].

6. References

D. G. Anderson and R. Barakat (1994). "Necessary and sufficient condition for a Mueller matrix to be derivable from a Jones matrix," J. Opt. Soc. Am. A. **11**, 2305-2319.

R. Barakat (1987). "Conditions for the physical realizability of polarization matrices characterizing passive systems," J. Mod. Opt. **34**, 1535-1544.

S. R. Cloude (1986). "Group theory and polarisation algebrea," Optik (Stuttgart) **75**, 26-36.

S. R. Cloude (1989). "Conditions for the physical realisability of matrix operators in polarimetry," in *Polarization Considerations for Optical Systems II*, R. A. Chipman, Editor, Proc. SPIE **1166**, 177-185.

J. J. Gil and E. Bernabeu (1985). "A depolarization criterion in Mueller matrices," Optica Acta **32**, 259-261.

J. J. Gil and E. Bernabeu (1987). "Obtainment of the polarizing and retardation parameters of a non-depolarizing optical system from the polar decomposition of its Mueller matrix," Optik (Stuttgart) **76**, 67-71.

S. S. Girgel' (1991). "Structure of the Mueller matrices of depolarized optical systems," Sov. Phys. Crystallogr. **36**, 890-891.

A. B. Kostinski (1992). "Depolarization criterion for incoherent scattering," Applied Optics **31**, 3506-3508.

Shih-Yau Lu (1995). "An interpretation of polarization matrices," Ph.D. Dissertation, Dept. of Physics, Univ. of Alabama at Huntsville, Huntsville, AL.

Shih-Yau Lu and Russell A. Chipman (1996). "Interpretation of Mueller matrices based on polar decomposition," J. Opt. Soc. Am. A. **13**, 1106-1113.

R. Simon (1981). "The connection between Mueller and Jones matrices of polarizaiton optics," Optics Communications **42**, 293-297.

C.V.M. van der Mee (1993). "An eigenvalue criterion for matrices transforming Stoke's parameters," J. Math. Phys. **34**, 5072-5088.

ON THE DESIGN OF
Nd:YAG, Nd:YVO$_4$ and CrTmHo:YAG
LASERS

Robert J. Hopkins
Graduate Student
Department of Physics

University of Central Florida/CREOL
4000 Central Florida Blvd.
Orlando, FL 32816-2900

ON THE DESIGN OF
Nd:YAG, Nd:YVO$_4$ and CrTmHo:YAG
LASERS

Robert J. Hopkins
Graduate Student
Department of Physics
University of Central Florida/CREOL

Abstract

The Nd:YAG and Nd:YVO4 lasers were operated at a wavelength of 1 micrometer. Different output couplers were compared and acousto-optic Q-switching was performed. Power measurements for cw operation were made. For the Q-switched experiments, a silicon photodetector was used to detect the pulses. In all of the preceding experiments, a Titanium Sapphire laser was used as a pump source. The CrTmHo:YAG was cooled to 77K and operated at 2.1 micrometers. It was pumped with a modulated diode laser and the output pulses were detected via a liquid nitrogen cooled InSb photodetector.

ON THE DESIGN OF
Nd:YAG, Nd:YVO4 and CrTmHo:YAG
LASERS

Robert J. Hopkins

Introduction

The CrTmHo:YAG laser operating at a wavelength of 2.1 micrometers is a candidate for a convenient means of producing eye safe laser radiation. One application is laser radar. It has previously been shown that this laser can be operated both at 77K and at room temperature.[1] The main goal of this summer research program was to replicate earlier results while developing the techniques required to build a near room temperature laser that could be useful for eye safe laser radar. The laser transitions near 1 micron of Nd:YAG and Nd:YVO4 have been shown to be viable transitions for laser radar applications.[2] As these lasers are well understood and therefore provide a good benchmark, these were the first to be attempted.

Methodology

With use of a Burleigh wavemeter the output of a Spectra Physics Ti:Sapphire was tuned to 808 nanometers. A cavity 40 centimeters long was constructed with a 20 centimeter radius of curvature output coupler and a 30 centimeter radius of curvature high reflector. A Perkin-Elmer spectrophotometer was used to determine the transmission of both mirrors at 1064 nanometers and also the transmission of the high reflector at 808 nanometers. A Nd:YAG crystal 4 millimeters in diameter and 5 millimeters long was placed 27 centimeters from the high reflector at the minimum beam waist of the cavity. A focusing lens was then placed an appropriate distance from the cavity to facilitate end-pumping the crystal. Transmission measurements were made to determine the power absorbed by the crystal. Several power measurements were taken on the Ti:Sapphire beam and then the beam was directed through the crystal and the transmitted power was measured. After proper alignment of the cavity and pump beam, the output power as a function of absorbed pump power was measured. Four different output couplers of the same radius but different transmissions were used.

Characteristics of two simple Q-switching techniques were explored next. A variable speed chopper was placed in the cavity to provide Q-spoiling. The average power was measured with a power meter and the pulse width was measured using a silicon detector. From these measurements, pulse energy and peak power were calculated. These measurements were made as a function of pulse repetition frequency with both crystals operating at the 1 micron wavelength. The chopper was replaced by an acousto-optic Q-switch and the same measurements were made at a frequency of 1 kilohertz for comparison. The Nd:YAG crystal was then replaced with a Nd:YVO4 cube 5 millimeters on a side and the previous experiment was repeated. This laser was Q-switched at a repetition rate as high as 5 kilohertz.

An absorption spectrum of the CrTmHo:YAG rod was acquired from a Perkin-Elmer spectrophotometer. Strong absorption was found at 799 nanometers. The rod of CrTmHo:YAG, 1 centimeter long and 5 millimeters in diameter, was placed in an EG&G dewar for cooling to the temperature of liquid nitrogen. A spherical cavity 5 centimeters long was built around the crystal. The 2.5 centimeter radius of curvature high reflector was also antireflection coated for 750 to 850 nanometers. The 2.5 centimeter output coupler was 99% reflecting at the 2.1 micrometer laser wavelength. The pump source was a 7 watt maximum Optopower 797 nanometer diode laser. It was fiber coupled to focusing optics with a focal length of about 1 inch. The wavelength of the diode was shifted to the absorption band at 799 nanometers via a thermoelectric cooler. The temperature of the diode was monitored with a preinstalled thermistor. The 2.1 micrometer radiation was detected with an EG&G Judson InSb photodetector, which was cooled to 77K. A 2.1 micrometer interference filter was taped over the detector in order to discern the 2.1 micrometer laser from the pump laser. The pump was modulated with a pulse generator to provide diode laser pulses with a width of 25 milliseconds and a frequency of 30 hertz. This allowed time for the crystal to cool between pulses.

Results

After construction of the 40 centimeter long cavity power measurements were made for the purposes of performance characterization with different output couplers. These results are plotted in Figure 1. A variable frequency chopper was placed on the output coupler side of the crystal in the cavity to provide Q-switching.

13-4

Nd:YAG
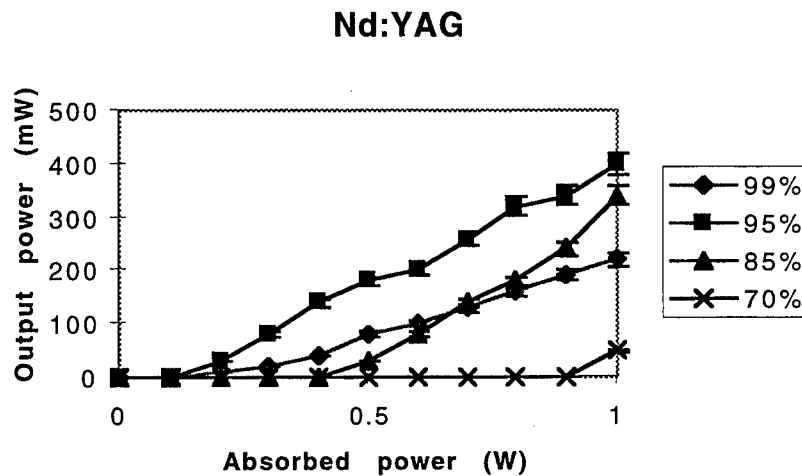
Figure 1

Figure 2 is a schematic representation of the laser cavity.

Nd:YAG/Nd:YVO4 setup
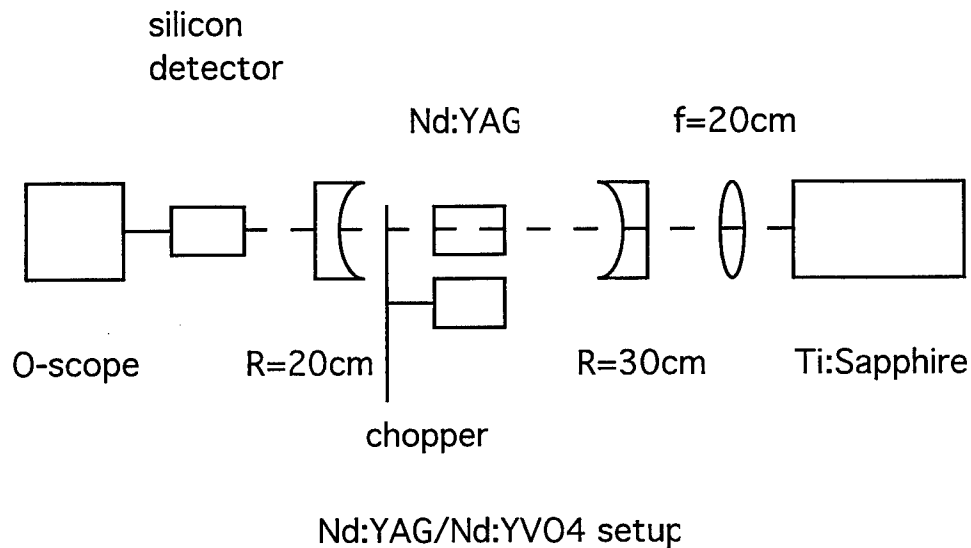
Figure 2

Average power and pulse width were measured with a power meter and a silicon photodetector, respectively.

Pulse energy and peak power were calculated. These values can be found in the table of Figure 3.

For the next experiment, the chopper was removed from the cavity and replaced with an acoustooptic Q-switch. It was operated at a frequency of 1 kilohertz. Average power was measured with the power meter at 20 milliwatts and the pulsewidth was 90 nanoseconds as measured with a photodetector and oscilloscope.

This yields an energy of 20 microjoules per pulse.

Copper Q-switch Nd:YAG

frequency (Hz)	average power (W)	pulsewidth (ns)	pulse energy (µJ)	peak power (W)
1000	.1	225	100	444
1500	.1	210	67	317
2000	.11	200	55	275
2500	.12	200	48	240
3000	.12	200	40	200
3500	.12	200	34	171

Figure 3

Consequently, the peak power level was about 220 watts. The Nd:YVO4 laser was Q-switched with the same acoustooptic Q-switch at a frequency of 5 kilohertz. The average power was 25 milliwatts and the pulsewidth was measured at 50 nanoseconds. Therefore the average energy per pulse was 5 microjoules and the peak power was 100 watts. A similar experiment using a chopper as a Q-switch was performed with the Nd:YVO4. The results are shown in the table of Figure 4.

chopper Q-switch Nd:YVO4

frequency (Hz)	average power (W)	pulsewidth (ns)	pulse energy (µJ)	peak power (W)
1000	.1	180	100	556
1500	.1	175	67	383
2000	.11	170	55	324
2500	.11	150	44	293
3000	.12	140	40	286
3500	.13	130	37	285

Figure 4

A portion of the absorption spectrum for CrTmHo:YAG is shown in Figure 5. The setup for the CrTmHo:YAG laser is depicted in Figure 6. Both the laser crystal and the detector were cooled with liquid nitrogen to 77K. The oscilloscope trace in Figure 7 shows the output of this laser.
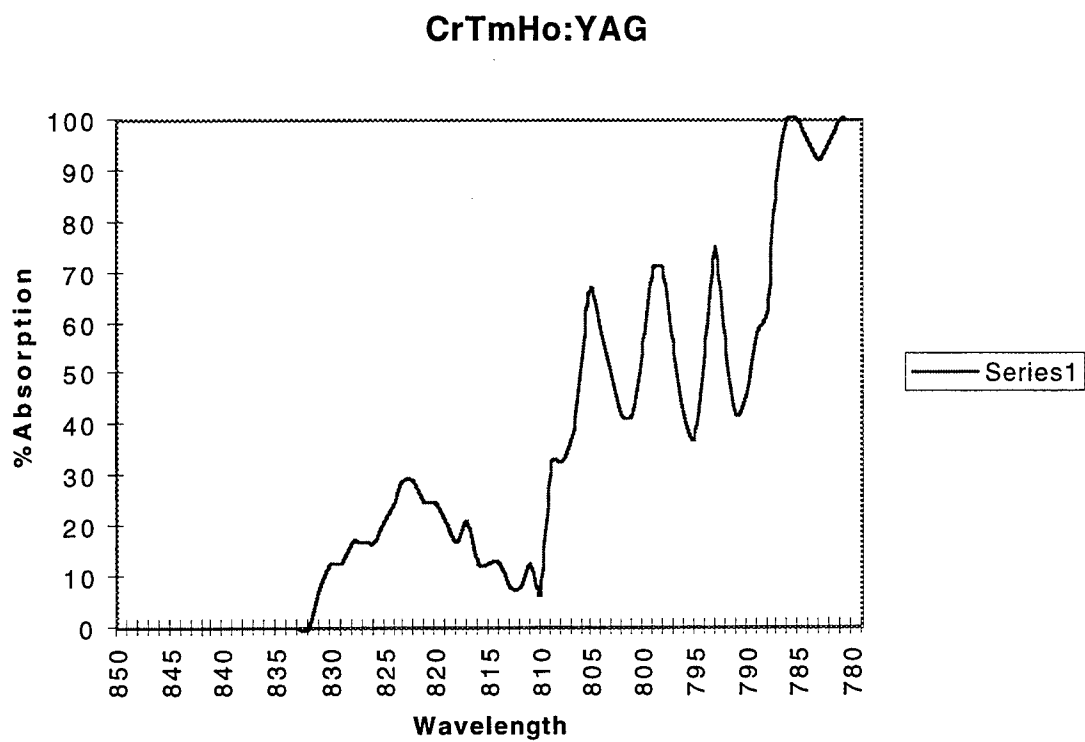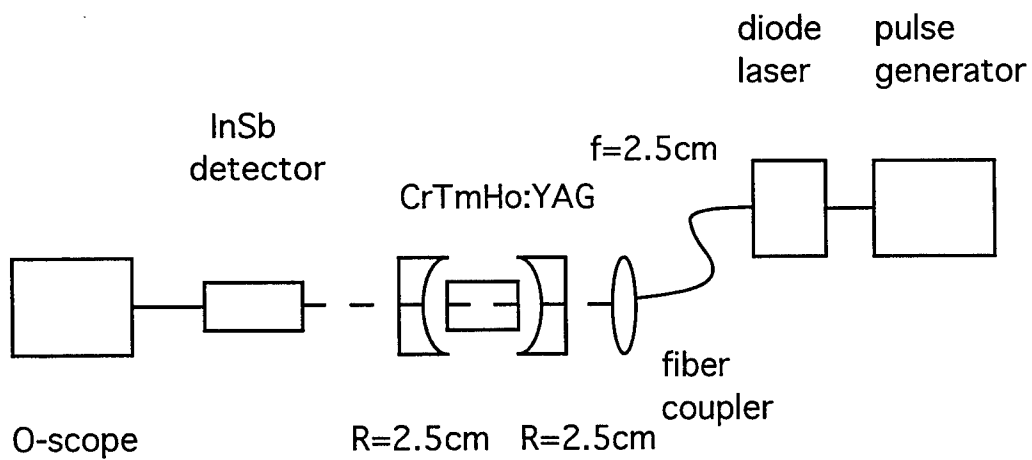
CrTmHo:YAG

Figure 5

CrTmHo:YAG setup

Figure 6

CrTmHo:YAG
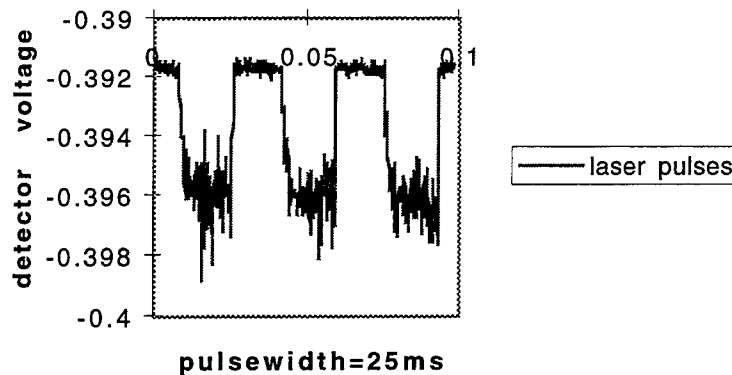
pulsewidth=25ms

Figure 7

Conclusion

Significant progress was made this summer. However, in order to make the CrTmHo:YAG laser applicable to laser radar it must be operated in a Q-switched configuration. Furthermore, the pulses must be short. A pulsewidth on the order of one to ten nanoseconds is necessary for our application. Previously, mode-locked pulses have been produced shorter than one nanosecond[5]. However, these pulses would have to be amplified in order to have the energy necessary for application to laser radar and these lasers are complicated devices. A shorter optical cavity would reduce the round trip time of the laser pulse within the cavity and thus shorten the pulse. Possibly due to problems with upconversion in holmium[6] the CrTmHo:YAG was not operated at room temperature.However, a shorter crystal perhaps 3 millimeters in length could suffice. Other crystal hosts should also be considered. Possibilities include BaY_2F_8 and $NaYF_4$. In these hosts there is a slower multiphonon relaxation rate and slower upconversion rates, therefore upconversion problems are minimized[6]. Further study should result in a viable system.

References

1. T. Y. Fan, G. Huber, R. L. Byer and P. Mitzscherlich, Opt. Lett. **12**, 678 (1987)

2. T. J. Kane, W. J. Kozlovsky, R. L. Byer and C. E. Byvik, Opt. Lett. **12**, 239 (1987)

3. V. G. Ostroumov, I. A. Scherbakov and A. I. Zagumennyi, OSA Advanced Solid-State Lasers, **15**, 64 (1993)

4. T. Y. Fan, G. Huber, R. L. Byer and P. Mitzscherlich, IEEE J. Quantum Electronics, **24**, 924 (1988)

5. F. Heine, E. Heumann and G. Huber, OSA Advanced Solid-State Lasers, **13**, 101 (1992)

6. K. M. Dinndorf, Ph.D. Dissertation, M.I.T. September 1993.

7. Tso Yee Fan and Robert L. Byer, IEEE J. Quantum Electronics, **24**, 895 (1988)

8. Y. Kalisky, S. R. Rotman, A. Brenier, C. Pedrini, G. Boulon, F. X. Hartmann, OSA Advanced Solid-State Lasers, **13**, 144 (1992)

AN AM1 STUDY OF BIPOLARONS IN DISCRETE CONJUGATED
MOLECULES WITH PENDENT ELECTRON WITHDRAWING GROUPS

David James Irvin
Graduate Student
Department of Chemistry

The University of Florida
P.O. Box 117200
Gainesville, Florida 32611

AN AM1 STUDY OF BIPOLARONS IN DISCRETE CONJUGATED

MOLECULES WITH PENDENT ELECTRON WITHDRAWING GROUPS

David James Irvin

Graduate Student

Department of Chemistry

The University of Florida

Abstract

Molecular modeling, using several programs, was used to predict the electronic properties of conductive polymers starting with discrete monomers, dimers and trimers and extrapolating to higher molecular weights. GAMESS, HyperChem, and GAUSSIAN92 were used to implement semiemperical AM1 calculations, and the results of the various programs were compared. The band gaps, torsion angles, and inter–ring bond lengths were calculated within GAMESS, and neutral and dicationic species were compared. The band gaps were calculated using S_1-S_1 and S_1-T_3 transitions, with the S_1-T_3 calculation yielding results closest to the experimental value of 1.9 eV. The rotational barriers were calculated at the AM1 level within GAUSSIAN92 for the neutral monomers only and are graphed as a function of angle. Using the results of the torsion angle and the inter-ring bond length calculations, the dication extends up to five rings.

AN AM1 STUDY OF BIPOLARONS IN DISCRETE CONJUGATED MOLECULES WITH PENDENT ELECTRON WITHDRAWING GROUPS

David James Irvin

Introduction

Research in electrically conductive polymers is highly active in the fields of chemistry, physics, material science, and electronics. With desirable properties such as electrochromism, electroluminescence, and photoluminescence, conductive polymers will not replace metals, but surpass them. The electron conduction in π-conjugated p-type doping polymers involves the migration of bipolarons, which are dicationic π-conjugated units that extend for two to seven rings. The formation, migration, and stability of these bipolarons effect the efficiency of electrical conduction.

In the recent literature, 3, 4–ethylenedioxythiophene (EDOT) has been shown to be electrochromic (indigo-blue to near colorless), to have a low monomer oxidation potential (1.0 V *vs.* Ag/Ag$^+$) and low polymer band gap (1.8 eV).{1} To further enhance the electrochromic properties, increase solubility, and tune the band gap, bis(3, 4–ethylenedioxythiophene)benzenes (BEB's) have been synthesized.{2} Palladium catalyzed cross-coupling reactions of substituted dibromobenzenes have been used in the synthesis of a wide variety of bisEDOTbenzenes, including alkyl, alkoxy, fluoro, hydroxy (protected), and acetate.{3} The disubstitution of dibromobenzenes with electron-donating and electron-withdrawing groups varies the electron density of the ring, influencing the electronic properties.

The field of computational chemistry has grown over the years in complexity and elegance to allow modeling of intricate systems, such as conducting polymers, or at least oligomers. Using programs such as GAUSSIAN92 and GAMESS, heats of formation, electron densities, and bond lengths, angles, and torsions can be calculated in times ranging from a few hours to a few days, depending on the complexity of the molecule and the modeling program. The relative energies can also be calculated for both singlet and triplet spin states. With calculations of the previous types, the band gaps of π-conjugated polymers can be accurately predicted.{4}

The impetus of this research was to determine if molecular modeling can predict the electronic properties of conductive polymers starting with discrete monomers, dimers and trimers and extrapolating to higher molecular weights. The model was tested on different programs to determine which of them calculates properties closest to the actual data received for two of the compounds modeled.

Computational Methods

A self-consistent field (SCF) restricted Hartree-Fock (RHF) approximation was used for closed shell molecules in neutral, dicationic singlet states. The SCF restricted open-shell Hartree-Fock (ROHF) approximation was used for open shell molecules in triplet states. The GAMESS{5} program was used for molecular orbital (MO) calculations. The heat of formation and equilibrium torsional angles were also calculated by *ab initio* methods employing the 6-31G* basis set.{6} Semiemperical calculations, such as Austin Model 1 or AM1, were implemented within GAMESS, HyperChem, and GAUSSIAN92. The results of the various programs were compared. The compounds calculated varied from 38 to 122 atoms and are shown in Figure # 1.
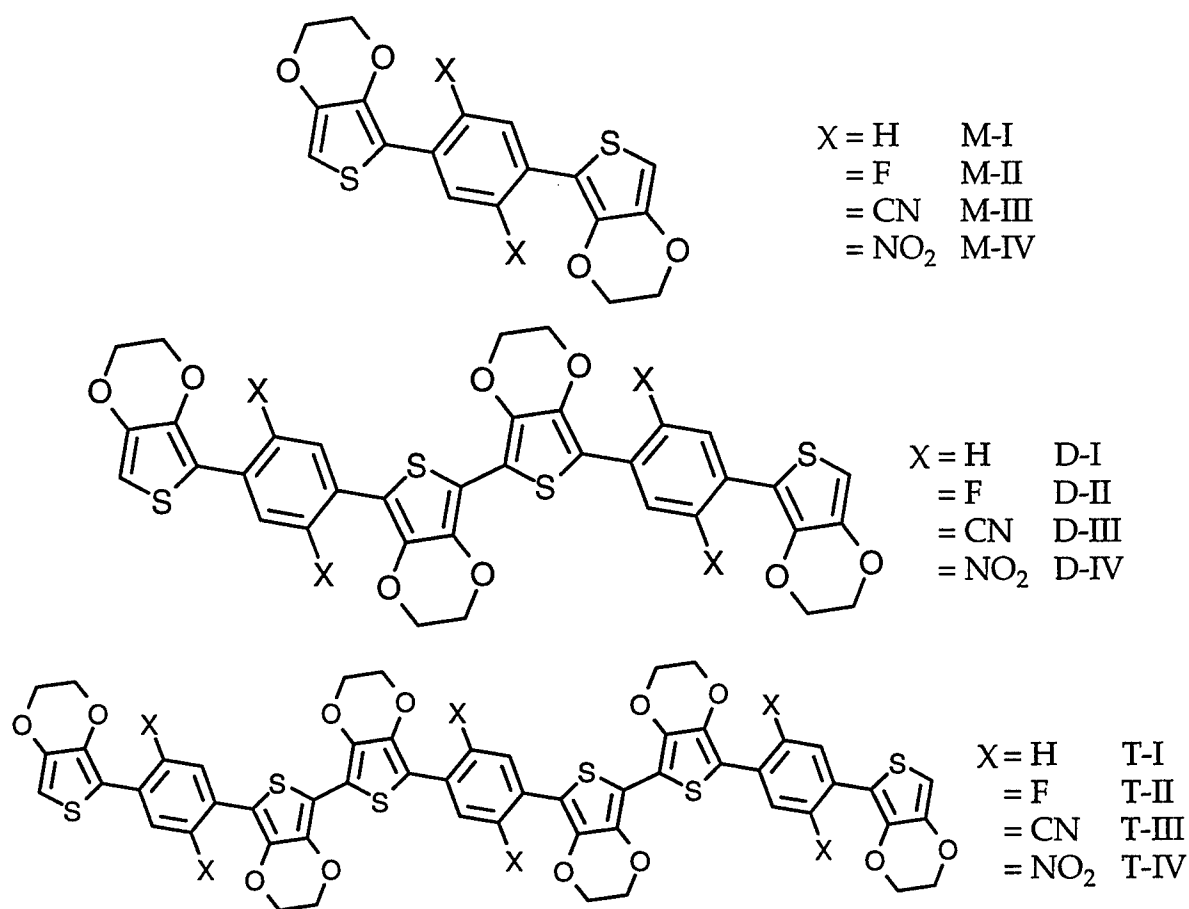
Figure # 1 : Structures of Compounds

The pendent groups hydrogen, fluorine, cyano, and nitro were used to vary electron density and steric bulk. The lattice relaxation or bipolaron stabilization energy was calculated by comparing the single point calculations of the geometrically optimized dication and of the dication in the neutral state geometry; the difference is the stabilization obtained by optimizing the geometry of the dication. The band gaps of the compounds were calculated by two methods. The first method involved calculating the Highest Occupied Molecular Orbital (HOMO) to Lowest Unoccupied Molecular Orbital (LUMO) gap using a singlet-singlet vertical excitation, and the second method was calculated for a singlet-triplet excitation. The inter-ring bond contractions were calculated from the difference in the optimized geometry of the dicationic species and

14-5

the optimized geometry of the neutral compound. The barrier of rotation was calculated for the monomers only using a matrix containing two dummy atoms; the AM1 calculation was implemented within GAUSSIAN92.

Results and Discussion

Since the methods of calculating the heat of formation vary from program to program, the absolute heat of formation can not be reliably determined. The change in the heat of formation, or ΔH_f, from the neutral to the dicationic species within each program is more reliable, and the results are shown in Tables I and II. Using AM1 within

Table I : Change in Heat of Formation and Torsional Angles By Program

Compound	G92			GAMESS			HyperChem		
	ΔH_f kcal/mol	Ø Z=0 (°)	Ø Z=2 (°)	ΔH_f kcal/mol	Ø Z=0 (°)	Ø Z=2 (°)	ΔH_f kcal/mol	Ø Z=0 (°)	Ø Z=2 (°)
M-I	421.6	5.9	7.8	418.8	17.1	1.0	418.8	17.5	1.4
M-II	422.3	10.2	12.2	430.2	43.8	1.4	x	38.4	x
M-III	436.2	8.9	10.4	434.8	23.2	1.2	x	x	2.0
M-IV	445.9	9.4	10.7	451.1	86.6	15.5	449.6	48.7	1.0

GAMESS, GAUSSIAN92 (G92), and HyperChem, ΔH_f's were calculated to within 7.9 kcal/mol of each other. This value is large considering that the same model was implemented by each program. In an experiment run to help elucidate the varying energies, single point calculations were run in GAMESS and in G92 for the GAMESS optimized matrix. A similar experiment was performed using a G92 optimized matrix. The Hartree-Fock (HF) energies calculated within G92 were higher for both matrices by 0.1174 Hartrees or 71.97 kcal/mol. The major reason for the above discrepancy is that G92 uses a Modified Neglect of Differential Overlap or MNDO basis set for sulfur while

Table II : Results from GAMESS AM1 Calculations

Compound	ΔHf (kcal/mol)	ΔHf[1] (eV)	Bipolaron Stablization Energy (eV)	Ø²[2] neutral (°)	Ø²[2] dication (°)	HOMO-LUMO Gap (eV) Z=0	HOMO-LUMO Gap (eV) Z=+2
M-I	418.82	18.16	1.0	17.1	0.9	7.48	5.25
M-II	430.22	18.66	1.5	43.8	1.4	7.82	5.13
M-III	434.81	18.85	1.0	23.2	1.2	7.33	5.12
M-IV	451.20	19.57	1.7	86.8	15.5	7.48	6.96/2.88
D-I	380.53	16.49	1.0	17.9	1.0	6.76	4.19
D-II	395.51	17.16	2.1	51.9	4.7	7.70	4.08
D-III	398.31	17.27	1.0	22.9	1.7	6.68	4.07
D-IV	428.64	18.59	1.6	53.2	45.6	7.28	4.31/2.80
T-I	367.56	15.94	1.0	18.1	4.0/2.7/0.4 4/3/2[3]	6.59	3.68
T-II	382.16	16.57	1.3	44.4	19.9/13.7/0.8 4/3/2[3]	6.93	3.28
T-III	385.53	16.72	1.2	26.6	5.9/3.1/0.7 4/3/2[3]	6.56	3.44
T-IV	399.86	17.34	1.2	39.7	20.4/16.1/7.3 4/3/2[3]	6.60	4.40/3.13

[1] Calculated by Hf dication -Hf neutral

[2] The average dihedral angle of all dioxyethylenethiophene/benzene rings

[3] The number of rings (phenyl and/or thienyl) from the center benzene

GAMESS uses a newly developed AM1 basis set. Within the G92 output it is stated that "...such mixtures of methods are very risky and have not been fully tested."

Within GAMESS, the ΔH_f ranged from 15.94 to 19.57 eV (Table II). The seemingly high energies are inflated since the calculations were done for the molecules in vacuum. The effect of solvent on the stabilization of the dications should be quiet high. The ΔH_f's decrease as the molecules increase in size, therefore increasing conjugation length. This follows the prediction that, with greater charge separation and a greater number of rotational and configurational modes, the ΔH_f will decrease. The ΔH_f was extrapolated for higher molecular weight polymers by graphing the ΔH_f as a function of one over the monomer unit (Figure # 2). The predicted ΔH_f for the polymer was between 14.8 and 16.8 eV depending upon the electron withdrawing substituent.

To determine the effect of conformational relaxation within each compound, the heat of formation was calculated for the dication in its optimized geometry. The value received was then compared to the energy calculated for the dication using the optimized geometry of the neutral compound. The geometrical rearrangement reduced the energy by approximately 1 eV (23.1 kcal/mol); these results are found in Table II under bipolaron stabilization energy. The stabilization energy was higher for the nitro containing compounds; this is due to possible charge interactions with the oxygens of the nitro, which are discussed later.

The difference in ring torsion angles between G92 and the other programs was significant. The change in the ring-to-ring torsion angles is also reported in Tables I and II. All of the neutral compounds were calculated in GAMESS and HyperChem to have inter–ring torsion angles of between 17° and 86°. A crystal structure of the M-I

14-8

compound confirmed the calculated result of 17°.{7} Only the nitro compounds' torsion angles change with an increase in molecular weight, from 86° for M-IV, to 40° for T-IV, with D-IV being intermediate at 53°.

All of the calculated torsion angles in G92 were within 4.3° of each other, while the angles were larger for GAMESS and HyperChem. The GAMESS and HyperChem output was in greater agreement with ΔH_f (<3.5 kcal/mol) and torsional angles (<5.4°) with the exception of the nitro compound. The difference in bond angles may be due to the MNDO basis set not accounting for the steric effects of the non-bonding electron pairs of the sulfur.
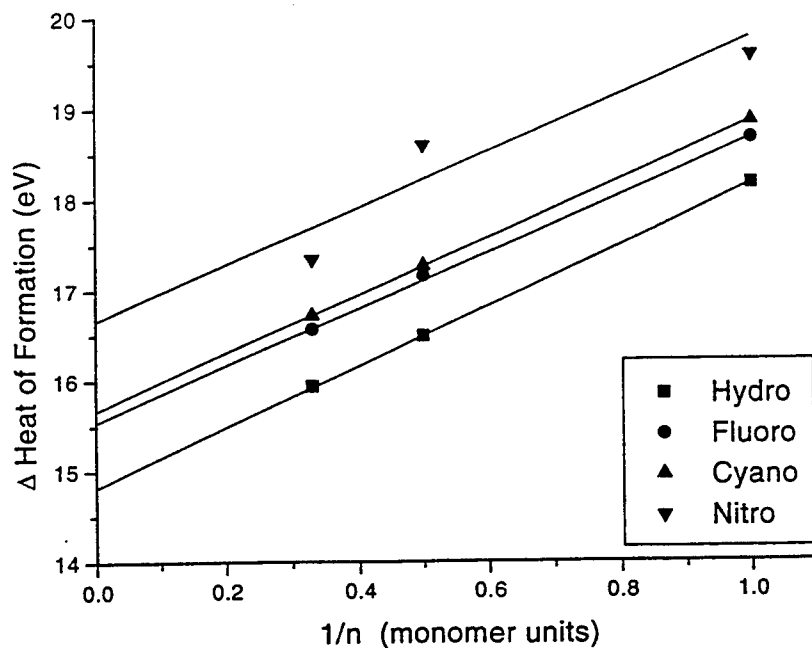
Figure # 2 : ΔH_f as a function of monomer units

14-9

As shown in Table II, all of the dicationic species exhibited a marked decrease in ring-to-ring torsion angles. All of the monomers showed a coplanar structure as dications, with the exception of M-IV. In the nitro monomer, one of the thienyl rings was coplanar, and the other maintained an angle of 30° out of plane, thus the calculated average was 15.5°. In an effort to determine why M-IV is unsymmetrical, the interatomic distances created by GAMESS were examined. One of the nitro oxygens was found to be 1.80Å from the thienyl sulfur in the dication. This gave credence to the structure proposed in Figure # 3, in which a six member heterocycle was formed on one side of the molecule. The bond distances of the other oxygens were too great (4.2Å and 2.4Å) to have strong bonding interactions, and both had a nitrogen to oxygen bond distance of 1.20Å. In addition, the oxygen-nitrogen bond length for the non-sulfur bonded oxygen was 1.17Å, while the interacting oxygen had a bond length of 1.30Å.

1= 1.80Å
2= 3.80Å
3= 4.20Å
4= 2.40Å
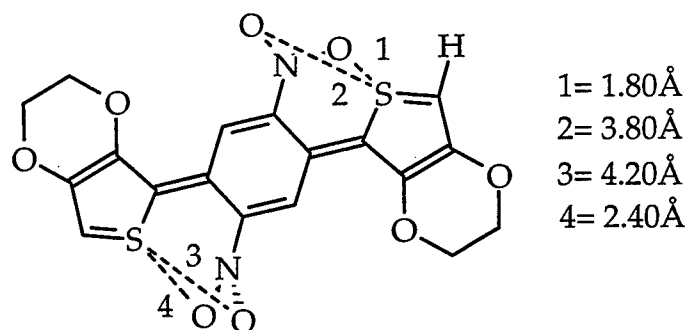
Figure # 3 : Calculated structure of M-IV

All of the dimers coalesced to a coplanar configuration over the entire six ring system, with exception of the nitro compound D-IV. In the trimers, all showed a coplanar structure for the five internal rings, and varying degrees of rotation, from 4° for T-I to 20°, for T-II and T-IV, in the external rings. For compounds containing pendent hydrogen, fluorine, and cyano groups, the neutral ring angles were the same for each

14-10

type as the chains were extended. The angles were 17°, 44°, and ~24°, for the hydro, fluoro, and cyano respectively, with a range of less than 4° for all. The ring-to-ring torsion of the fluoro compounds was unanticipated. Sterically, the fluorine is as small as hydrogen, but the ethylenedioxythiophene sulfur is dramatically repelled by the fluorine, resulting in the large torsion angle. The large torsion angle of 87° was expected for M-IV due to its steric bulk, but the decease in the angle in the extended chains was anomalous.

The difference in inter-ring bond lengths between neutral and dicationic species was correlated to the increase in ring-to-ring π-overlap. In Table III, the differences in inter-ring bond lengths are given. For M-I, M-II, and M-III, the decrease in bond length was symmetrical to the center, and all had approximately the same contraction, -0.08Å. The nitro monomer, M-IV, was unsymmetrical, with a net contraction approximately half that of the other monomers, -0.04Å. Compounds D-I, D-II, and D-III, all exhibited symmetrical bond contractions with the greatest contraction at the center of the molecule and decreasing contraction toward the termini. D-IV appears to concentrate the dicationic charge over four rings to one side of the molecule. Three of the bonds had similar contractions of approximately -0.06Å, and the other two bonds showed insignificant bond shortening.

In an effort to accurately predict the HOMO–LUMO gap, or band gap, in an undoped π-conjugated polymer, the HOMO–LUMO energy difference was calculated by two methods. The first was a singlet-singlet (S_0->S_1) transition, the results of which are found in Table II for the neutral as well as the dication states. The next method was a singlet-triplet (S_0->T_3) transition, the results of which are found in Figure #5 for only the neutral state. In the S_0->S_1 transition calculations, the band gap was determined to

be between 7.3 and 7.8 eV for all monomers. The dimers and trimers show the predicted decrease in band gap to 6.6 eV for all trimers except T-II, which was 6.9 eV. These values were very high compared with the spectroscopically determined band gaps of the polymers synthesized from M-I and M-II, which were 1.7 eV {2} and 1.9eV {3} respectively. In the S_0->T_3 transition , the monomers were calculated to have band gaps of between 2.1 and 2.3 eV. For compounds D-I, D-III, T-I, T-II and T-III, the band gap was determined to be ~1.75 eV. The nitro compounds started to decrease in band gap as the others, but the calculated value for T-IV was higher than for the monomer M-IV. The singlet-triplet excitation gave results much closer to that of the experimental data for the polymers of M-I and M-II.
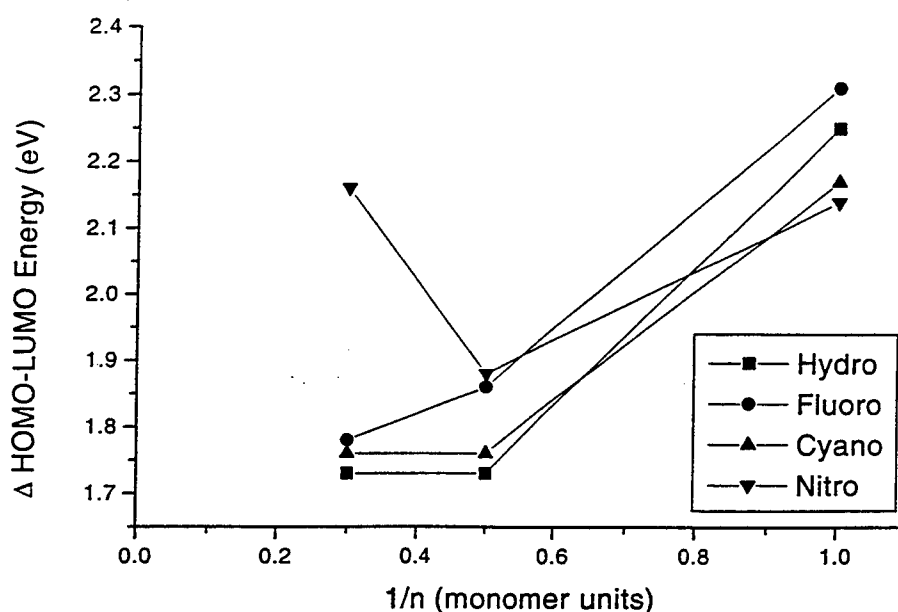
Figure # 4 : Δ Energy of HOMO-LUMO for Singlet-Triplet Transition

14-12

Table III : Change in Bond Lengths Between Rings (Å) (Z=2 - Z=0)

Compound	A	B	C	D
M-I	-0.075	-0.075	--	--
M-II	-0.080	-0.080	--	--
M-III	-0.076	-0.076	--	--
M-IV	-0.009	-0.030	--	--
D-I	-0.034	-0.055	-0.070	--
D-II	-0.036	-0.052	-0.073	--
D-III	-0.038	-0.046	-0.071	--
D-IV	-0.055/-0.002	-0.058/-0.007	-0.069	--
T-I	-0.009	-0.016	-0.048	-0.077
T-II	-0.008	-0.009	-0.046	-0.077
T-III	-0.011	-0.009	-0.048	-0.076
T-IV	-0.010	+0.003	-0.048	-0.071
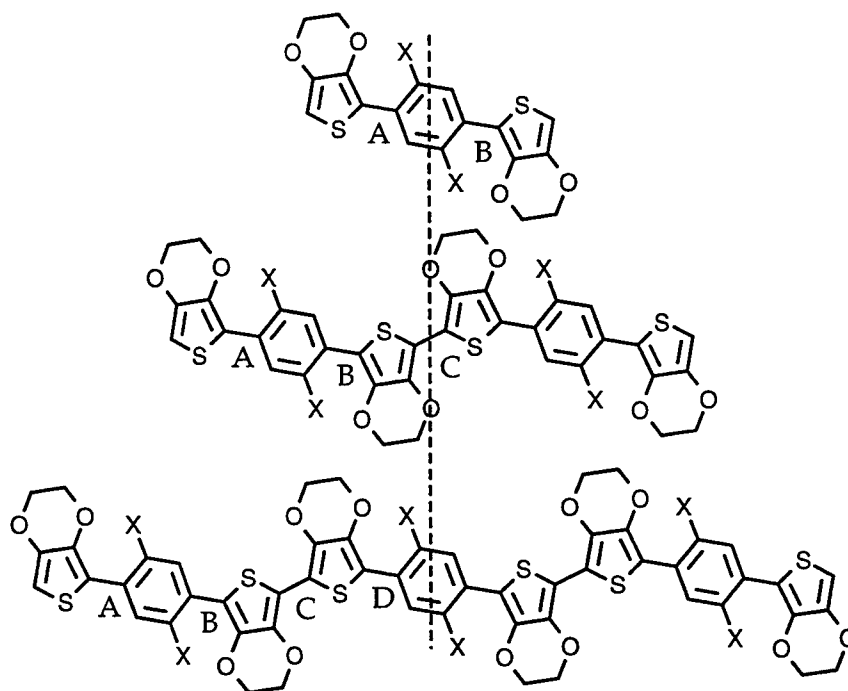
Figure #5 : Identification of Bonds for Table III

14-13

The rotational barriers were calculated for M-I through M-IV using z-matrices with two dummy atoms. The thienyl rings of the two EDOT rings were held stationary, and the internal benzene ring was rotated at 20° intervals. The lowest Hartree–Fock, or HF, energy calculated for the series of angles was used to normalize the energy. The value was halved, since the energy was calculated for both rings. The results are shown in Figure #6.
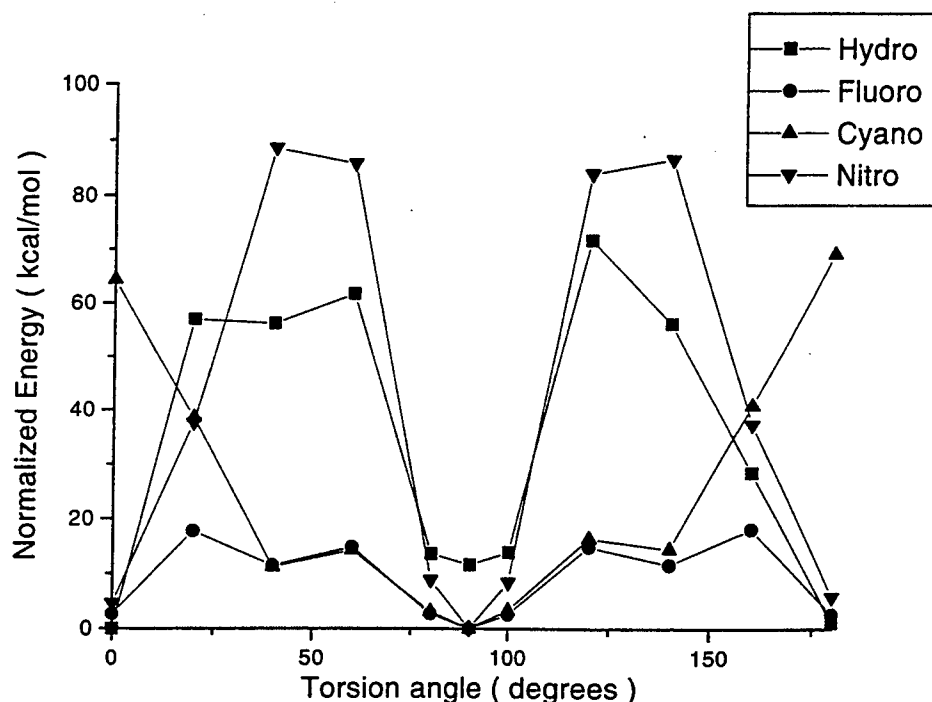
Figure # 6 : Normalized Rotational Energy for Monomers

Using the results of the torsion angle and the inter-ring bond length calculations discussed above, the apparent bipolaron length was determined. The monomers, with exception of M-IV, were calculated to extend the bipolaron over all three rings. The nitro compound is shown above. The dimers were calculated to spread the bipolaron

14-14

over 4 rings for the nitro-containing compound and over all six rings for the others. All of the trimers exhibited a bipolaron length of five rings in the center of the compound.

Conclusions

The band gap of an electrically conductive polymer can be accurately determined using computational methods. Best agreement with experimental values can be obtained by calculating the S_0->T_3 transition for discrete molecules and oligomers in GAMESS and then extrapolating the higher molecular weight. The bipolaron length was calculated to be five rings in higher oligomers.

References

{1} Pei, Q.; Zuccarello, G,; Ahlskog, M.; Inganäs, O. *Polymer*, **1994**, *35*, 1347.

{2} Reynolds, J. R.; Sotzing, G. A.; Sankaran, B.; Sapp, S. A.; Irvin, D. J.; Irvin, J. A.; Reddinger, J. L. *Polymer Preprints*, **1996**, *37(1)*, 135.

{3} Irvin, D. J.; Reynolds, J. R. *Polymer Preprints*, **1996**, 37(1), 532. and Irvin,D. J.; Welsh, D. M. unpublished results, **1996**.

{4} Beljonne, D.; Cornil, J.; Friend, R. H.; Janssen; R. A. J.; Brédas, J. L. *J. Am. Chem. Soc.*, **1996**, *118*, 6453.

{5} Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.*, **1993**, *14*, 1347.

{6} Hehre, W. J.; Radom, L.;Schleyer, P. v.R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*, Wiley, New York, **1986**, p80.

{7} Sotzing, G. A.; Reynolds, J. R.; Steel, P. J. *Chem. Mater.* **1996**, *8*, 882.

Numerical Simulation of Exploding Foil Initiators and Land Elements in PSpice

George W. Jarriel Jr.

Research Assistant

Department of Electrical Engineering

Auburn University

200 Broun Hall

Auburn, AL 36849-5201

George W. Jarriel Jr.

Research Assistant

Department of Electrical Engineering

Auburn University

Abstract

Numerical investigation of large arrays of exploding foil initiators (EFI), or slappers, is desirable

to minimize the number of design, fabrication and testing iterations during the development process.

Simple numerical models of the complete firesystem, including the the non-linear slapper element and the

land regions connecting the slappers to the rest of the circuit can be created with basic SPICE elements.

This model can then be used to parametrically analyze and optimize designs quickly and with enough

accuracy to guide developers. Creating an adequate model is dependent on the development of a accurate

model of the land regions for arrays with a large number of EFI's or large land regions.

1. Introduction

Exploding foil initiators such as the one shown in **Error! Reference source not found.** are used to initiate the detonation of explosives. A typical EFI method for fabricating slappers is to deposit a layer of conductor (such as copper or aluminum) on a substrate (i.e. mylar, alumina, silicon) and then etching away the excess copper to produce a geometry such at the one shown in Figure 1.

When large currents are passed through the device, the narrow "bridge" region is ohmically heated by the large current densities bunced inn the device. The metal foil passes quickly into a liquid state from the heating, then gasseous, and finally breaks down into a plasma due to the high temperature of the matereial and the high electrical fields corresponding to the currents in the foil. Once breakdown is achieved, the rapidly expanding plasma shears a section of the covering material off and propels it at great speed (~3 km/s[1]) at an explosive seperated from the assembly by a thin barrel. This sheared material is called the flyer, and when it strikes the explosive a shockwave (~3.2 GPa for a 15x15 mil flyer) is induced over a small region is produced, detonating the explosive.

Due to the extremely nonlinear behavior of slappers, as well as their relatively small size with respect to the land regions, traditional numerical methods (i.e. MOM, FEM) are difficult to apply and are economically prohibitive (i.e. they require an extreme amount of computer resources.) With this in mind, two methods have been developed to deal with these geometries; a finite elements code simulating only the land elements and a spice method coupled with a MOM code.

This report is concerned only with the spice method.

---

[1] The values given are meant to serve as examples only.

Simplified representation of 2x1 Slapper Array
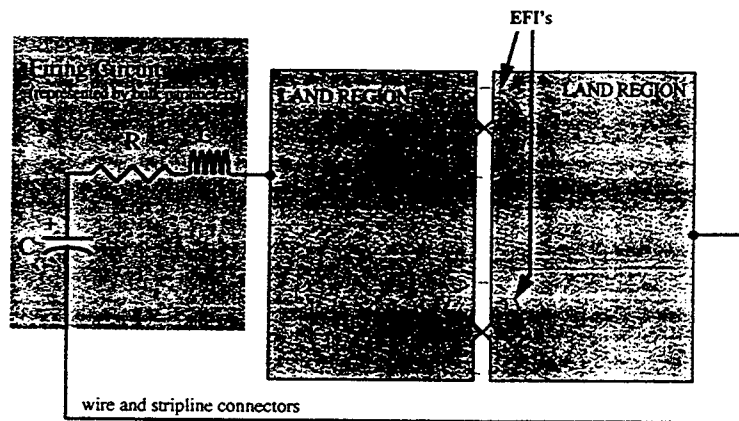
Figure 2 Simplified representation of 2x1 Slapper Array

The firing circuit (or the fireset) can be modeled by a lumped element LCR circuit. In figure 3 a typical fireset used for testing slappers is shown. The lumped element model of the circuit is shown in figure 2. The capacitance is a known value, and the numerical values for the other elements are generated using a simple ring-down test.

The resistance of the discrete regions are calculated using emperical data from the CRC manual [4. Weaver].

Capacitive distribution in the land regions of the EFI array are calculated using charge distribution profiles generated by a MOM code. Because the charge distribution is static in nature, the effects of the small bridges on the geometry are minimal, and thus can be ignored. This is not the case with the electrodynamic case where the EFI's have a greater effect on the system behavior.

The relationship between the charge distribution and the capacitance is given by the simple formula:

Key:
L - length of EFI
W- width of EFI
T - thickness of EFI
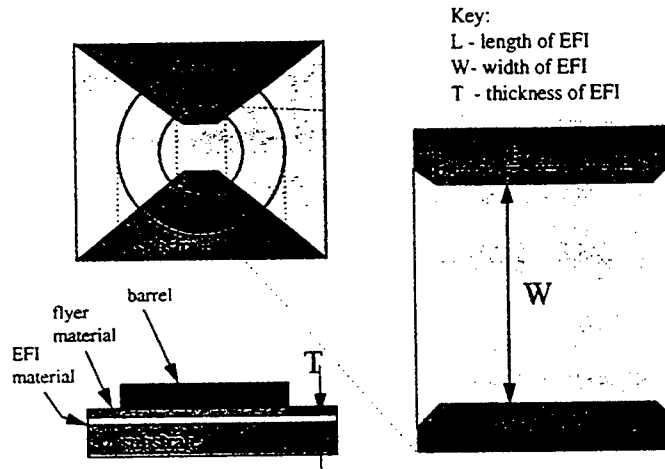
Figure 1- Typical EFI geometry.

2. Numerical Analysis: Modeling of Slappers and Land Regions With PSpice

As shown in Figure 2, an EFI array can be broken down into three distinct components for the purpose of numerical analysis: the firing circuit, the land regions and the slappers. Previous analysis using FEM techniques treated the entire circuit as a single unit and predicted the current distribution using first principles (Maxwell's equations). While this method is more desirable for preburst analysis, it is unacceptably slow for models more complex than a single slapper and does not adress the post burst behavior of the system. This is due to the difficulties involved in accurately modeling the intricate palsma interactions over the very short time of interest using a discrete technique.

$$q = \frac{1}{2} C \cdot V$$

where C is capacitance, q is the charge in the discrete region, and V is the the potential of the region.

Using the MOM code to determine the charge distribution, and then integrating over small discrete regions, an array of capacitance can be generated describing the capacitive distribution of the land regions.

Inductance does not play a large role in the dynamic current distribution for the geometries we are interested in at the present moment. The large stripline sections of the land regions will have small inductance, making them a secondary effect, and therefore negligable for the pupose of this research. Inductance in the fireset portion or the model is completely dominant, largely due to the capacitor and switch inductances.
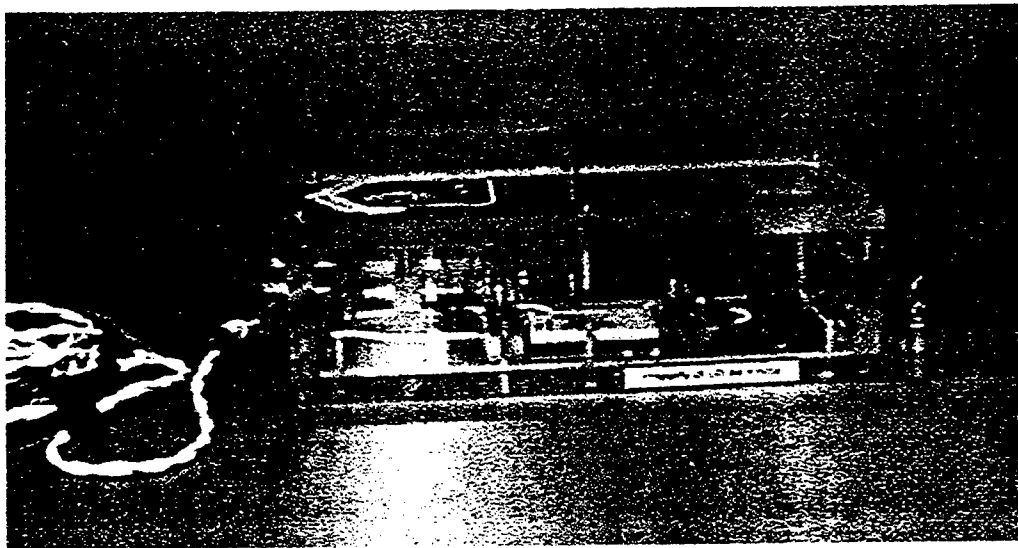
Figure 3 - Typical EFI fireset for initiation experimentation.

The flyer velocity can be determined from the burst current density of the slapper. Burst current is defined as the density of the current through the slapper at the time of slapper detonation. Statistical analysis of experimental results yields an empirical relationship given by:

$$E_{eg} = K \bullet J_b^n$$

where Jb is the burst current and K and n are constants determined experimentally. These constants change with variations in the slapper geometry.

3. Numercial Results

3.1 Single Slapper

In order to verify the results for a single slapper, a fire set such as the one shown in **figure ?** was employed to detonate several slappers. Before a fireset can be used to drive devices, it should first be characterized with a ring down test at several voltages in order for any slapper data to be meaningful. Specifications for the fireset are shown in Appendix A. These values will be used throughout the remainder of this report as the fireset values. Shown in **figure ?** is the equivalent circuit for the fireset using these values.

Conclusions

In order to develop a model sufficiently robust for what-if analysis, EFI modeling must incorporate emperical data. Attempts to analyze slappers from first principles suffer from two problems.

These methods are very slow, and generally speaking run into difficulties when the EFI reaches the plasma state, a state that is both unstable and extremely non-linear.

Incorporation of the emperical model allows for the use of Pspice as analysis tool, allowing for the analysis of the firing circuit, the EFI's and the land regions in a initiation array. The modeling of the land regions in spice is especially important, because it's incorporation into traditional modeling schemes would greatly increase the analysis time. As it has been shown, these land regions can be modeled easily with a simple RC array for the geometries considered here. Neglecting the inductive components in these arrays simplifies the problem and increases the resolution with which the problem can be looked at.

Further research will investigate the possible methods for considering land curvature, variations in slapper fabrication and materials other than copper land.

Acknowledgments:

Bibliography

1. Keith A. Thomas, Michael E. Baginski and Edward Shaeffer, A *Characterization of the Electrodynamics of Metals under the Action of Large Electric Currents.* International Journal on Electromagnetics, in preparation.

2. C. M. Furnberg, G. R. Peevy and S. G. Barnhart, *Slapper Detonator Modeling Using the Pspice Electrical Circuit Simulator.* Sandia National Laboratories Report, 1993.

3. Roger F. Harrington, *Electromagnetic Fields.* New York: McGraw-Hill, Inc, 1961.

4. Robert C. Weast, *Handbook of Chemistry and Physics: 56th Edition.* Cleveland, Ohio: CRC Press, 1976.

Appendix A- Fireset Specifications

Switch Type- Kryton

Stripline Dimensions-(Copper on polyemide)

    Copper:

        Length   10 inches

        Width   1 inch

        Thickness   1 mil

    Polyemide:

        Thickness   1 mil

        Capacitance   0.20 μF (3.5 kV maximum)

        Inductance   22.1 nH

        Resistance   55.0 mΩ
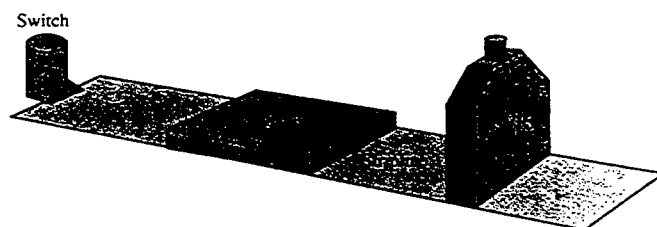
A STUDY OF WASTE REMOVAL PROCESSES
FOR A BARE BASE

Nicholas L. Jenkins
Graduate Student
Department of Civil Engineering

Georgia Tech
School of Civil and Environmental Engineering
Atlanta, GA 30332-0355

A STUDY OF WASTE REMOVAL PROCESSES
FOR A BARE BASE

Nicholas L. Jenkins
Graduate Student
Department of Civil Engineering
Georgia Tech

Abstract

There has been a concerted effort throughout the Department of Defense to develop appropriate methods for removing waste generated at a foreign military bare base. A bare base is an Air Force deployment of any number of airmen (this report will focus on 1100 airmen) that can be deployed anywhere in the world with only two requirements: a water source and operating surfaces. The objectives of this report are to create environmentally friendly techniques, similar to those used in the United States, systems that fit within the operational structure of a bare base contingency, and that conform to any international standards for waste treatment. The major focus of the research is to locate data and information that should help in creating the optimal(s) waste collection system.

A STUDY OF WASTE REMOVAL PROCESSES
FOR A BARE BASE

Nicholas L. Jenkins

INTRODUCTION

The purpose for the summer research was to assist in assessing the solid waste stream of an 1100 man deployable military force. The research was to look at quantifying and qualifying the waste streams generated on a bare base. Through numerous interviews (see interviews in reference section) with people directly involved with an Air Force bare base operation, it became readily apparent that one way to acquire accurate data was to actually take measurements from an existing bare base.

Fortunately, a new U.S. military bare base facility was currently being built from scratch in the deserts of Saudi Arabia. This presented a prime opportunity to collect some accurate data. A plan was devised to take these measurements, identify necessary equipment and procure the equipment. The plan could not be implemented because of budget and time constraints.

Cancellation of the planned visit to Saudi Arabia necessitated the development of other data collection methods. Three approaches were realized as possible data sources. The first was to look for actual data sources from a permanent air base such as Tyndall Air Force Base, FL. Interviews were conducted to locate personnel who had been assigned to bare bases who could possibly have knowledge and opinions relating to the existing waste collecting process. The second approach was to look for actual bare base data. The only information that was located was at the Army Corps of Engineers in Winchester, VA. The third approach was to examine the Silver Flag training site at Tyndall AFB.

Because the overall goal of the project is to improve upon the existing bare base waste collection system, numerous information on waste collection systems and techniques was collected and assessed. Vendors were contacted for information on their products and descriptive literature of collection systems were gathered from the base technical library. The main thrust of the research was to collect and analyze

data, therefore, solutions to problems recognized by this area of research will not be extensively discussed.

A brief description of what the author feels is the best waste collection system is provided in the conclusion section of this paper. It is based solely on information gathered during this short period of research and is susceptible to change once more information is gathered.

METHODOLOGY

AIR FORCE BASE DATA

Because the sponsoring lab (WL/FIVC-OL) is located in a remote site at Tyndall AFB, this base was the most convenient place to begin a waste stream data search. Interviews were conducted with the head of the base recycling program, the head of solid waste for AFCESA (Air Force Civil Engineering Support Agency), and bioenvironmental personnel in charge of hazardous waste collection and disposal. Each of these sources were able to provide consolidated historical data in their respective areas and were able to give further leads as to where more information might be located.

A day was spent on the recycling collection truck as the truck followed its weekly route to pick up recycling bins from the base housing units. The recycling director was interviewed on various aspects of the recycling program. The director stated it was difficult for the Air Force to make any money because of the recent drop in recycling prices (Munzel 1996). Additional funds are received from the sale of "expended 20 mm brass shells from live-fire F-15 runs over the gulf (Gulf Defender 3/3/95)". His prior combat experience led him to believe that the one major difference between a bare base and a permanent base is the lack of glass on a bare base because of its higher weight and susceptibility to breakage.

The head of solid waste at AFCESA is in charge of developing spreadsheet models for solid waste collections at bases around the country. He has collected detailed information on the quantities and types of waste that occur on one particular US air base. He is trying to develop a model that can be used for all bases to help determine the amount of waste generated at each base (Jacks 1996).

16-4

The head of hazardous waste disposal at Tyndall has been collecting data on all types of waste for the base. The data contains quantities for hazardous, non-hazardous, medical, base, and housing wastes. The system for data collection is relatively new and is continually being upgraded and improved.

LOGCAP Data

The data collected from the Army Corps of Engineers (COE) located in Winchester, VA was actually data compiled through LOGCAP, the Logistics Civil Augmentation Program. It is a U.S. Army program that uses a civilian contractor to perform engineering services and selected logistics to augment US forces during a military operation. The current contract was awarded to Brown and Root Services Corporation of Houston, Texas for $3.9 million in August 1992. The contract was for one year with four one-year options. The contract is now in the third one-year option.

The objective of the program is to set up, during peacetime, a worldwide management plan to provide logistical support (when required) to U.S. Army deployments. This rapid response contract is set up to provide basic needs, such as laundry, food, shelter, and sanitation, as well as provide maintenance and transportation services.

Brown and Root (B&R) has been required to support six contingencies during the LOGCAP contract. These include the following: Operation Restore Hope in Somalia, Operation Support Hope in Zaire, Operation Uphold Democracy in Haiti, Operation Vigilant Warrior in Southwest Asia, Operation Deny Flight centered out of Aviano Air Base in Italy, and Operation Joint Endeavor in Hungary. The latter two operations are directly related to the ongoing conflict in Bosnia.

One of (B&R) responsibilities during the conflict in Bosnia is refuse collection. They keep records on such things as: how much trash was collected, how much liquid waste was disposed of, how much fuel was used etc. This data is reported daily to B&R headquarters which is in turn sent to COE. The COE contact for the LOGCAP contract is the contracting officer at Winchester who helped give direction as to what information was actually available.

B&R supplied a summary sheet of events that occurred in Bosnia, Haiti, & Somalia. This included a sum total of solid, sanitary, and hazardous waste generated in each country that was handled by B&R.

In order to properly review this information it was necessary to restructure each data set. A data set consisted of 30 pages of information. Each page was the summary of one camp's information for the week in question. (To save time, the data sets that represented only one day were not used. Also there were actually 39 camps but the information received was for only 30.) The data contained numerous tasks such as the number of toilets, construction work in progress, power generation etc. It also contained items that can be related to actual waste generation. These included liquid and solid waste management, laundry bags washed, number of meals served, water and fuel delivery, hazardous materials collected, and population levels. The data for each category was entered in a spreadsheet. See Table 1. Data was not available for each item so some slots are left blank. For each category it was necessary to list the average for the week as well as for the particular day that the report was actually written. The reason for this was because there were inconsistencies in the reporting methods. Some numbers were given as the average for the week while others were just reported without any additional info. Also the units of measure was not consistent between camps so additional columns were created in which to place the unit of measure. If any hazardous material information was available it was also listed. As seen in Table 1, some of the information is not quantifiable because it is listed as number of dumps or loads for the solid waste or as mains (as in city water mains) for either the incoming or outgoing liquids. When the information was presented in this form it was not used in the calculations.

After the four data sets were entered, another spreadsheet was created that contained the data for each individual camp on a per person basis. See Table 2. This table shows the information for one base covering four different weeks. To remain consistent with the reported information, it was decided that liters would be used for the three liquids and cubic meters would be used for solid waste. The units of bags (laundry) and meals (food) were used consistently by all camp managers. Unit conversions were used to

convert all the data to the same set of units. (For example, if the data was given in gallons it was multiplied by 3.7854 to obtain liters. Similarly, data (not including solid waste) given in cubic meters was multiplied by 1000 to obtain liters.) Whenever possible, the average daily use for the week was used. If this was not available, then the reported daily use was listed. The average within each category is also shown.

To get an idea as to how much solid waste was generated per person per day, a third spreadsheet was created. See Table 3. This table lists the actual reported population and the solid waste generated for each camp that provided it for a week in May 1996. The solid waste numbers were converted to a per person basis in cubic meters. Since the density of the waste is unknown, the last three columns provide the range of densities commonly found in the municipal solid waste (MSW) stream for the United States. Seventy five cubic feet per ton is generally accepted as the average MSW density (Circeo, 1996). The numbers in each of these columns are given in pounds per person per day. At the bottom of the spreadsheet, the numbers for the camps with the highest and lowest quantities are shown as well as the average for all the camps. Thus, a serviceman generates trash on the order of 100 lbs/day.

According to one source (Steuteville 4/96), the 263 million Americans living in the United States generated a total of 326,709,000 tons of municipal solid waste in 1995. Thus the average American generated a total of 6.8 lbs/day:

$$327 \times 10^6 \text{ tons/year} * 2000 \text{ lbs/ton} * 1 \text{ year/365 days} / 265 * 10^6 \text{ people} = 6.8 \text{ lbs/person-day}$$

The average recycling rate was 27%, thus a total of 5.0 lbs/person-day (6.8*0.73) had to be disposed of through incineration (10%) or landfilling (63%).

The Bare Base Guide (Sverdrup 96) uses a planning factor of 4.0 lbs/day per soldier on a bare base. There are some differences between trash on a bare base and that of the average American. Very little glass is found on a bare base and recently much of the potable water has been bottled water causing a possible increase in the percentage of plastic in the waste stream. Other differences include the use of MREs on a bare base and the large amount of fast food trash generated in the US.

Taking a conservative estimate, the 6.8 lbs/person-day will be chosen as the baseline for municipal solid waste on a bare base. Why does the data from Bosnia give the impression that each soldier is generating roughly 100 lbs/person-day? The answer probably lies in the types of waste that are generated and measured. In the US, MSW does not normally include construction and industrial wastes. Currently, there is very little differentiation between types of waste on a bare base (Ridilla 96).

The problems with the data received from the COE were numerous. The camps were mostly supporting Army personnel not Air Force so the how this information would apply to an Air Force bare base is definitely in question. One basic difference between the Air Force and the Army is that the Air Force allows a higher quality of life for its servicemen. It is unknown what types of firepower the camp was supporting, i.e. a tank unit or a field artillery unit. Different units may generate different types of trash. The camp managers reporting the information were not aware of its importance so it was not consistently reported. Therefore, the relative exactness of the data is unknown. Only four useful data points were collected from the LOGCAP contract. The collection of more data would prove helpful in proving the validity of the data already obtained.

Medical Waste

Two means to estimate the amount of medical waste generated by a bare base have been found at this time. The first is the medical solid waste planning factors used by Air Combat Command in their Medical Deployment Support draft (essentially a bare base planning guide for medical support) (Air Combat Command, 1996). The second method is to look for any actual data kept by hospitals. (At the time of this writing, data has been promised but not yet received.)

Silver Flag

Silver Flag, located at Tyndall AFB, is the training arm of the RED HORSE Team. RED HORSE stands for Rapid Engineering Deployable Heavy Operations Repair Squadron Engineer. RED HORSE is

responsible for constructing the bare bases and building any major infrastructure improvements for existing air bases. The airmen that train at Silver Flag are members of RED HORSE, PRIME BEEF, PRIME RIBS, or are Civil Engineering personnel. PRIME BEEF stands for Base Emergency Engineering Force. PRIME RIBS stands for Readiness In Base Services. PRIME BEEF is essentially a smaller version of the RED HORSE team. A PRIME BEEF team consists of approximately 75 airmen where a RED HORSE team has approximately 400 airmen. PRIME RIBS is responsible for handling services on a base, such as laundry, kitchen, and janitorial services. RED HORSE is a completely self sufficient group. It has its own medical personnel, auto mechanics, service personnel, etc.

Silver Flag trains 160 airmen during each five day session. They are trained in how to build various tents, repair runways, hook up sewer lines, etc. It is a small microcosm of the operations performed on a bare base. For the purposes of trash generation, however, Silver Flag can not be assumed to generate the same proportions of trash as a bare base. Its unique mission is providing classroom and hands on construction training. Medical services are limited and airline shipment and packaging wastes are minimal. There are no flightline operations or maintenance, though there is some flightline repair. Permanent structures are not built and there are no combat or humanitarian missions.

Despite limited data availability at the Silver Flag site, it does offer one excellent trash generation source--the kitchen. As with a bare base kitchen, all meals are prepared based on base populations, since it is the sole eating facility. Meal proportions are identical for each airman, the same cleaning solutions are used, and the same types of cups, plates, and napkins require disposal.

Wastes from the kitchen include numerous organics, liquids, metals, dirt, etc. The majority of the food and kitchen supplies are shipped in cardboard boxes, with the individual contents packaged in cardboard (rice) or found in #10 tin cans (fruits). Some items are shipped in plastic bags (kool aid) or bottles (cleaning supplies). The silverware is plastic, the cups styrofoam, and the trays are made of cardboard. All floor scraps, including dirt and sand, are placed in the trash to keep the rodent population as low as possible. All metal pots and pans that are burned beyond use are thrown in the trash. Completely

used food generators are thrown away after any leftover gas is evaporated. Rags are reused continuously until they must be disposed of. All of the above mentioned materials are representative of both Silver Flag and a bare base facility.

RESULTS

The wastes on a bare base can be divided into six categories: municipal solid waste, medical, flightline (maintenance), hazardous (chemicals), munitions, and construction and demolition (C&D). Municipal waste can be approximated by using the average waste generated by Americans before recycling since recycling programs do not exist in a bare base contingency. Thus, personal waste of 6.8 pounds per person per day will be assumed as the baseline.

For medical waste, the only data available at this time is the draft planning factors used by Air Combat Command (ACC 1996). ACC recommends the following hospital sizes given the number of personnel it supports. The approximate solid waste generated for each hospital are also given.

Hospital Size	# of personnel	Solid Waste Generated (lbs/day)
14 bed	500 - 1400	6,000
25 bed	1500 - 3000	11,600
50 bed	3000 - 5000	18,500

There will also be much larger medical centers set up at individual camps that will handle large number of casualties derived during conflicts. Since this study is only looking at a typical 1100 man bare base, the 6,000 lbs/day of solid waste generated by a 14 bed ATH will be used as a baseline for medical waste and the possibilities of having a larger central medical center will not be considered.

It is not possible to forecast the amount of waste generated from either munitions or C&D because these are directly related to the mission of the bare base. Hazardous waste and waste generated on the flightline side of the bare base are equally difficult to measure and there is no available data on these types of wastes.

CONCLUSIONS

Of all the waste generated on a bare base, only medical and personal waste need to be removed on a regular basis. This is dictated because organic material found in this type of trash attracts disease carrying rodents and insects and has a tendency to release unpleasant aromas. Fortunately, these types of bare base waste can be reasonably approximated. The other types of trash found on a bare base can be safely stored for extended periods of time (due to their inorganic nature). Thus, instead of trying to quantify and qualify the exact nature of the waste, it would be more appropriate to design a system that has the capacity to handle both the wastes that need immediate disposal and that also has additional capacity to handle other portions of the waste stream, as needed.

Although there are many different methods to handling the solid waste stream, there are only four reasonable alternatives available that do not use the existing host country's waste removal methods. These are the following:

1. The present practice which includes open pit burning, unlined and uncovered landfills and sanitary landfills (Silver Flag 1996).
2. Building a lined landfill similar to those found in the US. It would be covered daily with dirt in order to prevent rodents and odors to get out of control (Robinson 1986).
3. Build an incinerator.
4. Build a plasma torch.

There are various ways to augment these alternatives such as adding balers, compactors, and shredders but these do not change the overall methods to which they would be applied. On a bare base, it is an accepted practice to construct landfills wherever it is convenient (Silver Flag 1996). Therefore, when the first two methods are used, the trash will be removed to a place that is generally most convenient. When that area

is no longer useable, another convenient area will be chosen. The point of this is to show that the amount of waste generated is not important for these two methods.

This line of thinking also applies to the last two methods. Both the incinerator and the plasma torch can be operated at less than maximum operating power if necessary. Similarly, it would be futile to design a system that does exactly what is needed and nothing else. Therefore the system would most likely have a factor of safety involved which would allow for any large peaks in the waste stream.

Since the purpose of this study is to find ways that can handle the bare base waste stream in a manner consistent with the on-going environmental movement, the present methods of removal are not the choice of the future. Thus the other three methods must be compared to determine which is the best fit. To build a lined landfill requires an extensive amount of time and labor (and thus money) to build and maintain and has the potential to be a perpetual liability. An incinerator requires possibly a large initial capital outlay and also has a problem with ash residue. The only apparent drawback with the plasma torch is its possible large initial capital cost. It does not appear to have a residue problem since the output form the plasma torch is a geologically stable glassified rock. Given these three alternatives then, it appears that the plasma torch is the best alternative at this time.

Based on this conclusion, the following is the list of assumptions and operating parameters used to determine the approximate size for a generic plasma torch system:

1. MSW for a bare base is of the same density as the average MSW generated in the US.

2. Twenty hours of operation are allowed per day

3. No cogeneration of power is considered since it is not commercially available at this time.

4. Peak requirements call for a factor of safety of 1.5.

5. A bare base generates the same amount of MSW per person-day as the US.

6. Operations efficiency of 75%.

7. Medical waste is generated at the rate of 6000 lbs per day.

Using very simple calculations with the assumptions above, the size plasma torch needed would be approximately 190 kW. This is (1100 people * 6.8 lbs/day-person + 6000 lbs) * 0.75 * 1.5 / 2000 lbs/ton * 500 kW-hr/ton / 20 hr/day ≈ 190 kW. Any of the above assumptions could be grossly incorrect but for the purposes of this phase of the project they are as good as any.

Camp	Date	Current Population	Projected Population	Solid Waste (cu. m.)	avg	Liquid Waste (cu. m.)	u	avg	Hazardous Materials (barrels)	Laundry (bags)	avg	Food (meals)	avg	Water Delivery (liters)	u	avg	Fuel Delivery (liters)	u	avg	Notes
1	5/5/96	2700	2250	245		27	cm			284	284	3415		126000	-	77000	6089	-	6089	
2	5/5/96	1500	1700	248		12	cm			164	164			106800	-		3729	-		
3	5/5/96	1700	1700	207	207	16	cm			245	245			128000	-	128000	6362	-	6362	
4	5/5/96	1715	2150	100	100	2400	g	2400		188	188	1676	1676	mains			7388	-	7388	
5	5/5/96	650	723	30	30	mains				72	72	942	942	mains			350	g	350	
6	5/5/96	239	250	90	90	mains	g			42	41	250		250000	-	250000	1127	-	1127	
7	5/5/96	150	154	30	30	27000	-	27000		11	11			mains						
8	5/5/96	330 day	250 day	60	60	13500	-	13500						10000	-	10000				
9	5/5/96	955	960	57	57	37919	-	37919	8 for week	110	110	1195	1195	36000	-	36000	3741	-	3741	
10	5/5/96	1206	1225	65	65	26000	-	26000		138		1020	1020	137500	-	137500	1765	-	1765	
11	5/5/96	910	760	45	45	27000	-	27000		78	78	1414	1414	63644	-	63644	2579	-	2579	
12	5/5/96	73	99	2 dumps	2 dumps	mains				10	10			mains						
13	5/5/96	650	1027	44	44	34348	-	34348		75	75	1222	1222	75202	-	75202	1734	-	1734	
14	5/5/96	138	185	2 dumps	2 dumps	mains				10	10			mains						
15	5/5/96	106	131	2 dumps	2 dumps	mains			*	17	17			mains						*4 drums, 20-5gal cans so far
16	5/5/96	147	160	10	10	15250	-	15250		36	36			14400	-	14400	0	-		
17	5/5/96	286	Work Site	44	44	5810	-	5810		21	21			15840	-	15840	98	-	98	
18	5/5/96	160	Work Site			mains								mains			0			
19	5/5/96	393	446	8 dumps	8 dumps	3000	g	3000		66	66	479	479	10500	g	10500	253	g	253	
20	5/5/96	1150	1100	252	252	5000	g	5000	**	165	165	850	850	86750	g	86750	1066	g	1066	** 210 kg/wk of medical laundry 30-55gal drums
21	5/5/96	1353	1559			6000	g	6000	20/wk exp	220		726	726	32150	g	32150	210	g	210	
22	5/5/96																			
23	5/5/96	330	380	60	60	168	-			36	36	496	496	47000	-	47000	1708	-	1708	of fuel picked up
24	5/5/96		Work Site	1 load		mains			1/wk exp					180000	-	180000	0			
25	5/5/96	522	516	8	8	28000	-	28000		61	61	961	961	39500	-	39500	3865	-	3865	
26	5/5/96	354	379	16	16	11355	-	11355		54	54	592	592	149000	g	149000	924	g	924	
27	5/5/96	872	642	6	6	27300	-	27300		94	94	1402	1402	114660	-	114660	608	-	608	
28	5/5/96	1148	1177	70	70	28000	g	28000		1237		2322	2322	70000	-	70000	2500	-	2500	
29	5/5/96	643	627	35		21	cm			89	89	1311	1311	72100	-	72100	1942	-	1942	
30	5/5/96		Work Site																	

Table 1. Bosnia Data -- Week Ending May 5, 1996

				Data is in units per person per day					
Camp #	Date	Current Population	Projected Population	Solid Waste (cu. m.)	Liquid Waste (liters)	Laundry (bags)	Food (meals)	Water Delivery (liters)	Fuel Delivery (liters)
1	3/31/96	2700	2250	0.119	4.6	0.096	1.09	23.0	3.93
1	4/6/96	2700	2250	0.124	15.6	0.092	1.03	26.9	3.89
1	4/28/96	2700	2250	0.161	17.4	0.097	1.12	28.5	2.75
1	5/5/96	2700	2250	0.091	10.0	0.105	1.26	28.5	2.26
Avg		2700	2250	0.124	11.9	0.098	1.13	26.7	3.20

Table 2. Data for One Bosnia Camp for Four Weeks

Date: May 5, 1996

			Lbs/capita-day Density of Trash		
Population	Waste cubic meters	Waste/person cubic meters	40 CF/TON	75 CF/TON	100 CF/TON
2700	245	0.09	160	85	64
1500	248	0.17	292	156	117
1700	207	0.12	215	115	86
1715	100	0.06	103	55	41
650	30	0.05	81	43	33
239	90	0.38	665	355	266
150	30	0.20	353	188	141
330	60	0.18	321	171	128
955	57	0.06	105	56	42
1206	65	0.05	95	51	38
910	45	0.05	87	47	35
650	44	0.07	120	64	48
147	10	0.07	120	64	48
286	44	0.15	272	145	109
1150	252	0.22	387	206	155
330	60	0.18	321	171	128
522	8	0.02	27	14	11
354	16	0.05	80	43	32
872	6	0.01	12	6	5
1148	70	0.06	108	57	43
643	35	0.05	96	51	38
Totals: 18157	1722				
	Low	0.01	12	6	5
	High	0.38	665	355	266
	Average	0.09	167	89	67

Table 3. Solid Waste Measurements

REFERENCES

Bare Base

Air Combat Command, Draft of <u>Air Combat Command Instruction 41-151: Medical Deployment Support</u>, September 1, 1996.

Army Corps of Engineers, LOGCAP Situation Reports from Bosnia, 1996.

Brown & Root Services Corporation, Telefax on LOGCAP <u>Event Waste History</u>, July 7, 1996.

Sverdrup Technology, <u>Bare Base Infrastructure Road Map: Report on the Bare Base Study Phases I and II</u>, March 26, 1996.

Liquid Waste

Burrows, W.D., Schmidt, M.O., Carnevale, R.M. and Schaub, S.A. "Nonpotable Reuse: Development of Health Criteria and Technologies for Shower Water Recycle", *Water Science and Technology*, 1991, **24**(9), 81-88.

Calspan Corp. A study concerned with the identification of the characteristics of various wastewater streams generated from military bare bases. There was no available title. Sponsored by USAMERDC

Lawrence Livermore National Laboratory. Press Release: "Laboratory Deionization Technology May Herald Low Cost Sea Water Desalination and Waste Water Treatment", December 20, 1994.

Plasma Torch Technology

Camacho, S.L. "Harnessing Artificial Lightning," *The World & I*, December 1991, 310-317.

Camacho, S.L. "Plasma Pyrolysis of Hydrocarbon Wastes", Proceedings of Plasma for Industry and Environment: A Technical Conference, September 1990. BNCE and Wadham College, Oxford.

Mayne, P.W. and Mulholland, J.A. (Editors). Proceedings of the International Symposium on Environmental Technologies: *Plasma Systems and Applications*, Atlanta, GA, October 1995. Georgia Institute of Technology and University of Bordeaux I, France.

Plasma Torch Ventures. "Mixed Waste Disposal/Recovery Plant, 250 Tons/Day Module" (A brochure).

Savannah River Technology Center. "A Successful Solution for Waste Disposal" (A brochure). US Department of Energy.

Scientific American, Editors of. "The Ultimate Incinerators," *Scientific American* September 1995, 180.

Tuan, C.Y. and Dass, W.C. "A Feasibility Study of Using Plasma Arc Technology to Create Pavements." Report completed for Wright Laboratory, Tyndall AFB, June, 1996.

Weston Inc. "Preliminary Review of Project Reasonableness Plasma Arc Waste Processing System." Prepared for Bob Silverman of The Winter Group of Companies, July 14, 1993.

Recycling

Gulf Defender. "Reaping the Rewards of Recycling" 3/3/95, p 4-5.
Means, R.S. *The Change Notice*. 2nd Quarter 1995.

Steel Recycling Institute. *The Recycling MAGNET*. Various Issues.

Steel Recycling Institute. *The Southeastern MAGNET*. Various Issues.

Steel Recycling Institute. *SteeLinks*. Various Issues

Solid Waste

Barlaz, M.A. et. al. Developing a Life-Cycle Inventory of Municipal Solid Waste Management Alternatives -- System Description. Research Triangle Institute, NC, March 24, 1995.

Franklin Associates, Ltd. Characterization of Municipal Solid Waste in the United States, 1960 to 2000 (Update 1988). US Environmental Protection Agency. 1988.

Franklin Associates, Ltd. Characterization of Municipal Solid Waste in the United States: 1992 Update. US Environmental Protection Agency. July 1992.

International Ash Working Group. An International Perspective on Characterization and Management of Residues from Municipal Solid Waste Incineration. December 1994.

Kilgroe, J.D. (Chair). Solid Waste Management: Thermal Treatment & Waste-to-Energy Technologies, Proceedings of an International Specialty Conference, Washington, D.C., April 1995. Air & Waste Management Association.

Navy, United States Department of the. Draft Environmental Impact Statement: Disposal of US Navy Shipboard Solid Waste, April 1996.

Robinson, W.D. (Editor). The Solid Waste Handbook: A Practical Guide, John Wiley & Sons, NY, 1986.

Rogoff, M.J. "Waste Joins the Net Set." *World Wastes*, May 1996, 36-44.

Steuteville, R. "The State of Garbage in America, Part I." *BioCycle*, April 1995, 54-63.

Steuteville, R. "The State of Garbage in America, Part I." *BioCycle*, April 1996, 54-61.

Waste-to-Energy

AIMS Coalition, "Waste-to-Energy: Making a Clean Energy Source Cleaner" (A brochure), 1994. AIMS is made up of ASME, IWSA, MWMA, and SWANA.

Berenyi, E.B. and Gould, R.N. "Municipal Waste Combustion in 1993," *Waste Age*, November 1993, 51-56.

Charles, M.A. "New Trends in Waste-to-Energy," *Waste Age*, November 1993, 59-62.

Freeman, D. "Waste-to-Energy Fights The Fires of Opposition," *World Wastes*, January, 1994, 24-28.

Hathaway, S.A. Recovery of Energy from Solid Waste at Army Installations. August 1977. Construction Engineering Research Laboratory.

Malloy, M.G. and McAdams, C.L. "The US & International Municipal Waste Combustion Industry," *Waste Age*, November 1994, 89-98.

Ubaldi, R.A. "Using Waste-to-Energy to Repower Utility Electrical Generating Stations," *Energy Engineering*, 1994, **91**(5), 66-74.

Wilson, D.G., Saba, E.F., Nuwayhid, R.Y. and Hamrin, D. "A Waste-to-Energy Recycling Plant for Beirut," *Proc. Instn Mech. Engrs, Part A*, 1995, **209**, 63-70.

Interviews

Capt. Bainger, Tyndall AFB medical services, September 6, 1996.

Velice Bet-Sayad, Senior Systems Engineer for Sverdrup Technology, Eglin AFB, July 18, 1996.

Dr. Lou Circeo, Director of Construction Research Center at Georgia Tech, July 3, 1996.

Sgt. Dewar, Tyndall AFB hazardous material collection, July 22, 1996.

Mr. Bob Gruber, Contracting Officer for LOGCAP, Army Corp. of Engineers, Winchester VA, July 24-25, 1996.

Gary Jacks, Head of solid waste for AFCESA, Tyndall AFB, July 31, 1996.

Jim Munzel, Manager of Tyndall AFB Recycling Program, July 23, 1996.

Captain Pete Ridilla, Chief of Environmental & Safety for 823rd Red Horse Squadron, Hurlburt Field, July & August 1996.

John Tagg, Associate Principal Engineer for Sverdrup Technology, Eglin AFB, July 18, 1996.

Major Yatkus et.al., Commander of Silver Flag Training Site, July 1996.

SOL-GEL-DERIVED COATINGS FOR SPATIALLY CONTINUOUS PRESSURE MAPPING

Jeffrey D. Jordan
Graduate Research Associate
Department of Chemistry

State University of New York at Buffalo
Natural Sciences Complex
Buffalo, New York 14260-3000

SOL-GEL-DERIVED COATINGS FOR SPATIALLY CONTINUOUS PRESSURE MAPPING

Jeffrey D. Jordan
Graduate Research Associate
Department of Chemistry
State University of New York at Buffalo

Abstract

The accurate determination of pressure fields over aerodynamic test surfaces is critical to state-of-the-art aerospace development. Computational fluid dynamic (CFD) modeling of unsteady-flow phenomena requires experimental data for development and validation. Surface-measurement schemes using pressure-sensitive and temperature-sensitive paint technology have exhibited potential for the generation of spatially continuous pressure maps of test surfaces. The development of pressure-sensitive coatings has recently been complemented by sol-gel technology. The key advantages of sol-gel-derived composites over conventional paint-binder materials are two-fold. First, sol-gel-derived coatings are characterized by high thermal stability, and are thus good candidates for extreme-temperature applications. Second, novel thin-film architectures have been designed which provide a convenient means to produce thin, uniform films, *and* optimize the sensitivity of these composite coatings for the specific pressure and temperature regimes of interest. Together, these features reflect the potential of sol-gel-derived materials to broaden the applicability of pressure-sensitive paint technology. This approach provides a powerful tool for aerodynamic research that will aid the development and evaluation of CFD models and benefit both the commercial and the military aircraft industries in the United States.

SOL-GEL-DERIVED COATINGS FOR SPATIALLY CONTINUOUS PRESSURE MAPPING

Jeffrey D. Jordan

Introduction

The accurate determination of pressure fields and forces over aerodynamic test surfaces is critical to state-of-the-art aerospace development. Computational fluid dynamic (CFD) modeling of unsteady-flow phenomena requires experimental data for development and validation. Currently, a reliable, accurate means of measuring unsteady-flow phenomena on aerodynamic test surfaces is not available. Conventional pressure measurements are based on discrete pressure taps and electronically-scanned transducers. This approach provides accurate pressure information; however, several key drawbacks exist. First, the very nature of the pressure taps limits their ability to provide accurate pressure information at discrete points. Second, integration of the pressure taps on a test surface is time consuming, labor intensive and expensive. Third, the aerodynamics of the test surface are compromised by the pressure taps, which introduces the potential for bias in the measurement. Finally, pressure taps are subject to high failure rates, particularly for turbomachinery or extreme-temperature applications.

Pressure-sensitive paint (PSP) technology has emerged as an attractive alternative for determination of both static and transient surface-pressure fields for aerodynamic applications.[1-4] In these schemes, pressure sensitivity is based on the quenching of photoluminescent molecules dispersed in a polymer film which is sprayed onto a test surface. The optical sampling method allows both pressure and temperature measurements to be made with the same instrumentation, allowing their parallel acquisition. Figure 1 outlines the pressure-measurement concept. In this scheme, a photoluminescent species (A) absorbs light of appropriate energy $h\upsilon$ and is promoted to an excited electronic state (A*). The excited-state molecules then return to the ground-electronic state either by some radiative-decay process involving the emission of a photon (fluorescence), or by a non-radiative relaxation pathway. Collisions with ground-state oxygen molecules (O_2) can also result in additional non-radiative decay of the luminescent species to the ground state. This phenomenon, known as dynamic quenching, forms the basis for luminescent-based pressure measurements.
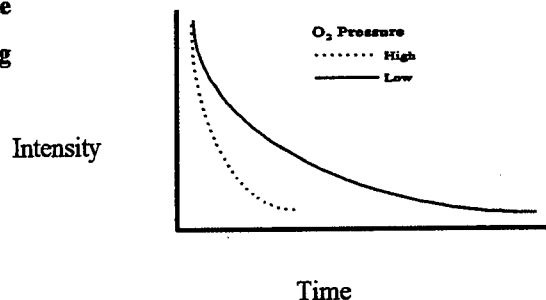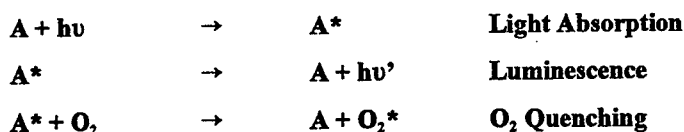
Quenching processes are modeled by some form of the Stern-Volmer relationship.[5] In its simplest form, the equation is given by

$$\tau_0/\tau = I_0/I = 1 + K_{sv}P_{O2} \tag{1}$$

where I and τ represent the luminescence intensity and excited-state lifetime, at a given P_{O2}, respectively, and the 0 subscript denotes measurement in the absence of quencher (vacuum). In application, a modified form of the Stern-Volmer equation is typically used which replaces the vacuum calibration with a reference standard. In this approach, the luminescence intensity at a reference pressure level, I_{ref}, is divided by the luminescence intensity at some test condition, I, over the region of interest. Calibration of this intensity ratio (I_{ref}/I) provides the surface pressure. The intensities are generally sampled over the surface of interest by a detector array such as a CCD camera. The output of the array can be visually represented as an image, with the luminescence corresponding to

a gray or false-color scale.

Photoluminescence and Quenching Processes

A + hυ	→	A*	Light Absorption
A*	→	A + hυ'	Luminescence
A* + O₂	→	A + O₂*	O₂ Quenching

$$A + h\upsilon \rightarrow A^* \qquad \text{Light Absorption}$$
$$A^* \rightarrow A + h\upsilon' \qquad \text{Luminescence}$$
$$A^* + O_2 \rightarrow A + O_2^* \qquad O_2\ \text{Quenching}$$

Intensity

Time

Temporal Behavior of Photoluminescence

$$d[A^*]/dt = -k_t\,[A^*]$$
$$k_t = k_1 + k_q[O_2]$$
$$[A^*]_t/[A^*]_0 = \exp(-k_t t)$$

Figure 1. Pressure-measurement concept based on O₂ quenching of photoluminescent probe molecules.

Statement of the Problem

Pressure-sensitive paint technology has emerged as a powerful alternative for determination of both static and transient surface-pressure fields for aerodynamic applications.[1-4] However, the major impediment to the routine use of this scheme is the lack of a general methodology for the production of pressure-sensitive coatings exhibiting the desired performance characteristics. The ideal protocol would provide a convenient means to produce robust coatings exhibiting optimal: 1) pressure sensitivity; 2) temporal response; and 3) thermal-, mechanical- and photostability for specific test conditions of interest.

Current efforts at Wright Laboratories focus on the development of pressure-sensitive coatings for the measurement of aerodynamic flow fields on turbomachinery. For this application, the test article is the suction surface of a compressor rotor under standard operating conditions. Specifically, the pressure gradients encountered at rotor rotational speeds exceeding 13 000 rpm are between 0.5 and 1.5 atm (7-21 psi), with a temperature maximum of 150 °C at the rotor-blade edge. For this application, the ideal pressure-sensitive coating would exhibit: 1) maximum sensitivity between 0.5 and 1.5 atm, 2) sub-ms temporal response (to probe transient phenomena faster than 1 kHz), 3) photostability, and 4) thermal stability.

The majority of pressure sensitive paints presented in the literature incorporate silicone-based binders.[1-4]

17-4

In general, silicone-based binders exhibit good pressure sensitivity in the 0 to 0.5 atm regime; however, this is at the expense of high-end pressure sensitivity (> 1 atm). Moreover, silicone-based binders are subject to temperature instability because the luminescent probe can sublime from the polymer support or the polymer matrix can reorganize. As a result, silicone-based coatings are not suitable for high-temperature or high-pressure applications. Thus, the widespread use of PSP technology remains contingent on the development of a methodology which allows the production of stable composite films exhibiting *tunable* performance characteristics.

Methodology

Recent advances in sol-gel technology offer a pathway to develop pressure- and temperature-sensitive paints. Sol-gel-derived composites offer several advantages over conventional paint binder materials. First, sol-gel-derived coatings are characterized by high thermal stability, and are thus good candidates for extreme-temperature applications. Second, a deposition technique which allows the formation of thin, uniform films has recently been developed,[6,7] and has been shown to produce optically transparent, robust composite coatings which possess molecular scale (< 0.1 μm) surface roughness. Control over the film thickness and surface roughness serves to effect changes in pressure and temperature sensitivity, improve temporal response, and avoid problems associated with laminar- and turbulent-flow boundary layers. Further, it is important to note that these sol-gel-derived coatings can be easily removed from the test surface by mild base treatment, leaving the test surface clean. Third, novel thin-film architectures have been designed which provide a convenient means to *optimize* the sensitivity of these composite coatings for the specific pressure and temperature regimes of interest.[8] Finally, sol-gel-derived thin films exhibit fast temporal response to pressure changes, allowing transient phenomena (> 1 kHz) to be quantified.

Obtaining pressure-sensitive coatings which exhibit temperature stability requires control over two processes. The first is the retention of the luminescent probe within the coating. Sublimation of probe molecules results in probe loss and distribution within the polymer matrix, which invariably modulates the photophysics of the luminophore and hence the performance characteristics of the coating. This effect can be reduced by using a luminescent probe with a relatively high vapor pressure, or via covalent attachment schemes if necessary. Second, thermally-induced reorganization of the polymeric *composite* can also result in modulation of the system dynamic response. Therefore, the generation of an accurate, spatially-continuous pressure map of a test surface requires a composite film/coating whose performance characteristics do not change over the test time frame. Recently, Bright and coworkers,[6,9,10] have demonstrated that stable sol-gel-derived chemical and biosensor platforms can be produced using a layered thin-film architecture. In this approach, the molecular recognition chemistry is "sandwiched" between upper and lower sol-gel films, which impart greater stability and increased loading capabilities. In the current work, sol-gel-derived films were used to produce pressure-sensitive coatings exhibiting greater temperature stability, and *tunable* sensitivity. The composite architecture is illustrated in Figure 2.

Upper sol-gel film
Active Layer
Lower sol-gel film
Primary aluminum substrate
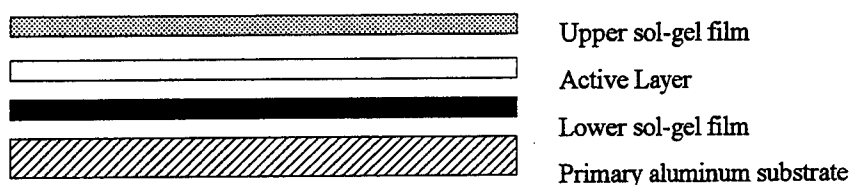
Figure 2. Sol-gel-derived sandwich architecture.

In this scheme, the lower sol-gel film, or shield layer, serves as protective barrier between the aluminum substrate and the active layer. This serves to: eliminate substrate-induced quenching of the fluorophore dispersed within the active layer, and reduce scattering of the excitation radiation. Together, the sol-gel layers act to increase the temperature stability of the active layer by perturbing thermally-induced reorganization of the polymer matrix, and probe sublimation. Further, the O_2 diffusivity through the upper sol-gel film can be affected simply by changing the thickness (number of coats applied) and/or porosity (chemistry of the sol-gel process). Thus, this approach provides *tunability* of the pressure sensitivity of the composite coating. Due to the proprietary nature of the preparation protocols used to produce the sol-gel-derived thin films, the full details will not be disclosed at this time.

The instrumental apparatus setup used for the characterization of the pressure-sensitive coatings is outlined in Figure 3.
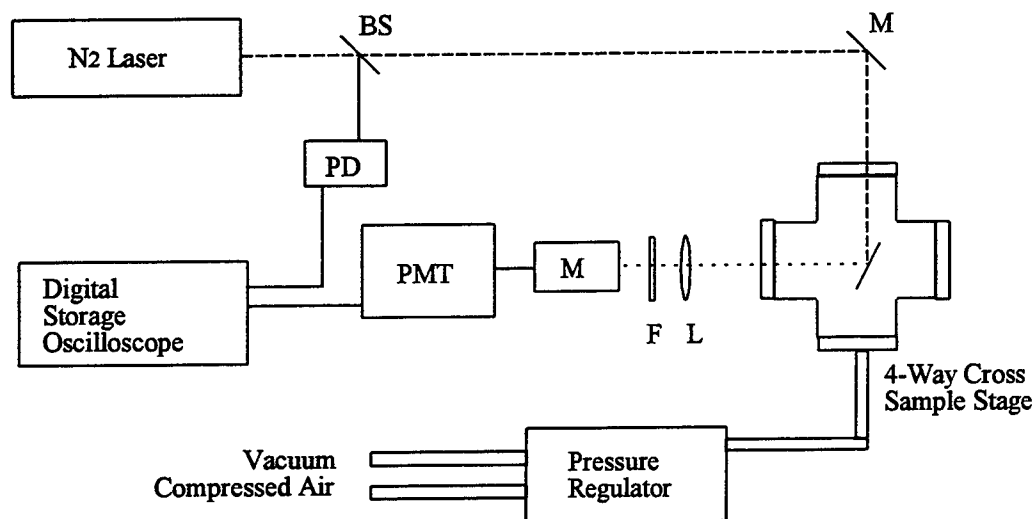
Figure 3. **Schematic of the time-resolved spectrometer used for characterization of pressure-sensitive coatings.**

A fraction of the pulsed (maximum repetition rate 20 Hz) 337 nm output from a nitrogen (N$_2$) laser is directed to a fast photodiode which triggers a digital storage oscilloscope. The remainder of the beam is directed into a four-way cross in which the sample is fixed at a 25 ° angle with respect to the incident beam. Emission from the test surface is collected, filtered and focused onto the entrance slit of a monochromator (2 nm bandpass) before impinging on a fast-wired photomultiplier tube (PMT). Time-resolved decay traces are then collected using a digital storage oscilloscope. The atmosphere within the four-way cross is controlled by a pressure regulation device which uses vacuum and compressed air in order to maintain a given cell pressure. The entire system is under computer control, and data acquisition is performed using software written in house. In operation, the cell pressure is incremented according to predetermined start and stop values, and time-resolved decay traces are collected at each point and stored. From this data, the excited-state fluorescence lifetime and total intensity under the decay curve is determined, and the resulting Stern-Volmer plots produced. Typical output from the post-process software is shown in Figure 4.
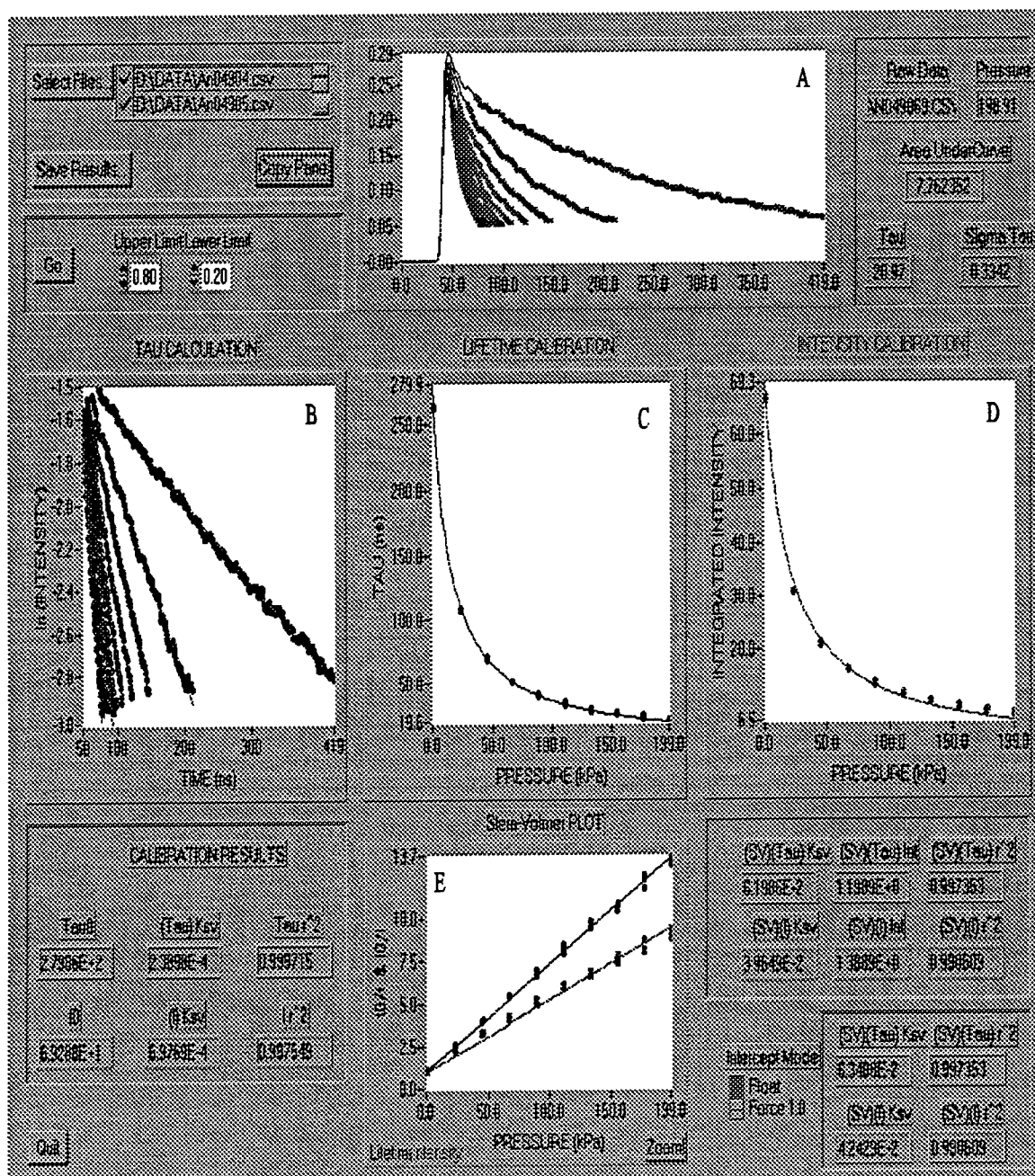
Figure 4. Typical output from post-process software.

The software developed performs simultaneous analysis on replicate data sets (n ≥ 3 in all cases). Panel A shows raw intensity decay traces for sol-gel-derived sandwich films over the pressure range 1 to 199 kPa (~ 0.01 to 2

atm). The fluorescence lifetimes at each pressure are determined from the slope of the ln(Intensity) vs time plots (Panel B), and are then used to produce a calibration curve (lifetime (tau) vs pressure, Panel C). Similarly, the intensity-based calibration (Panel D) is performed by plotting the total fluorescence intensity (*area* under the decay curve) as a function of air pressure. From these data, Stern-Volmer plots (Panel E) with the statistically recovered constants are generated. Finally, all data and graphs are saved using the root filename and identity-specific extensions, facilitating additional post processing.

Results

The ideal paint characteristics for the current application include: 1) thermal stability, 2) photostability, 3) maximum sensitivity between 0.5 and 1.5 atm, and 4) sub-ms temporal response (to probe transient phenomena faster than 1 kHz). The ability of the sol-gel layers to impart greater stability to the silicone-based pressure-sensitive paint was tested by preparing a series of sol-gel-derived sandwich coatings (SGS) and silicone-based paint (SP) samples. These samples were allowed to dry under ambient conditions for 15 h prior to performing the initial pressure calibration (*vide supra*). The films were then subjected to a 150 °C environment (in air, convection oven) for 90 min or 64 h. Following each heat treatment, the samples were allowed to cool to room temperature and recalibrated. The results of the lifetime calibrations (n = 3) for the PSP and SGS coatings are given in Figures 5 and 6, respectively. Three main points deserve mention. First, the pressure sensitivity (slope of the lifetime vs pressure curve) of the PSP decreased markedly as a function of heat treatment, culminating in negligible response after 64 h. In contrast, the sensitivity of the SGS remained stable under the same conditions. Second, the excited-state fluorescence lifetime of the probe in PSP decreased on heat treatment, while subtle, yet clear increases were observed for the SGS coatings. Third, a decrease in total intensity of nearly 70 % was observed for the PSP samples, indicating probe sublimation from the silicone binder. No intensity decrease was detected for the sol-gel-derived coatings. Together, these results indicate that the sol-gel sandwich architecture imparts temperature stability to the active layer by minimizing probe sublimation (no intensity decreases) and polymer reorganization (stable response sensitivity). Further, the increase in lifetime observed for the SGS coatings suggest that temperature pre-treatment can produce composites with stable pressure-sensitivity, independent of temperature history. This is particularly attractive for high-pressure applications, where severe temperature gradient fields are often encountered.
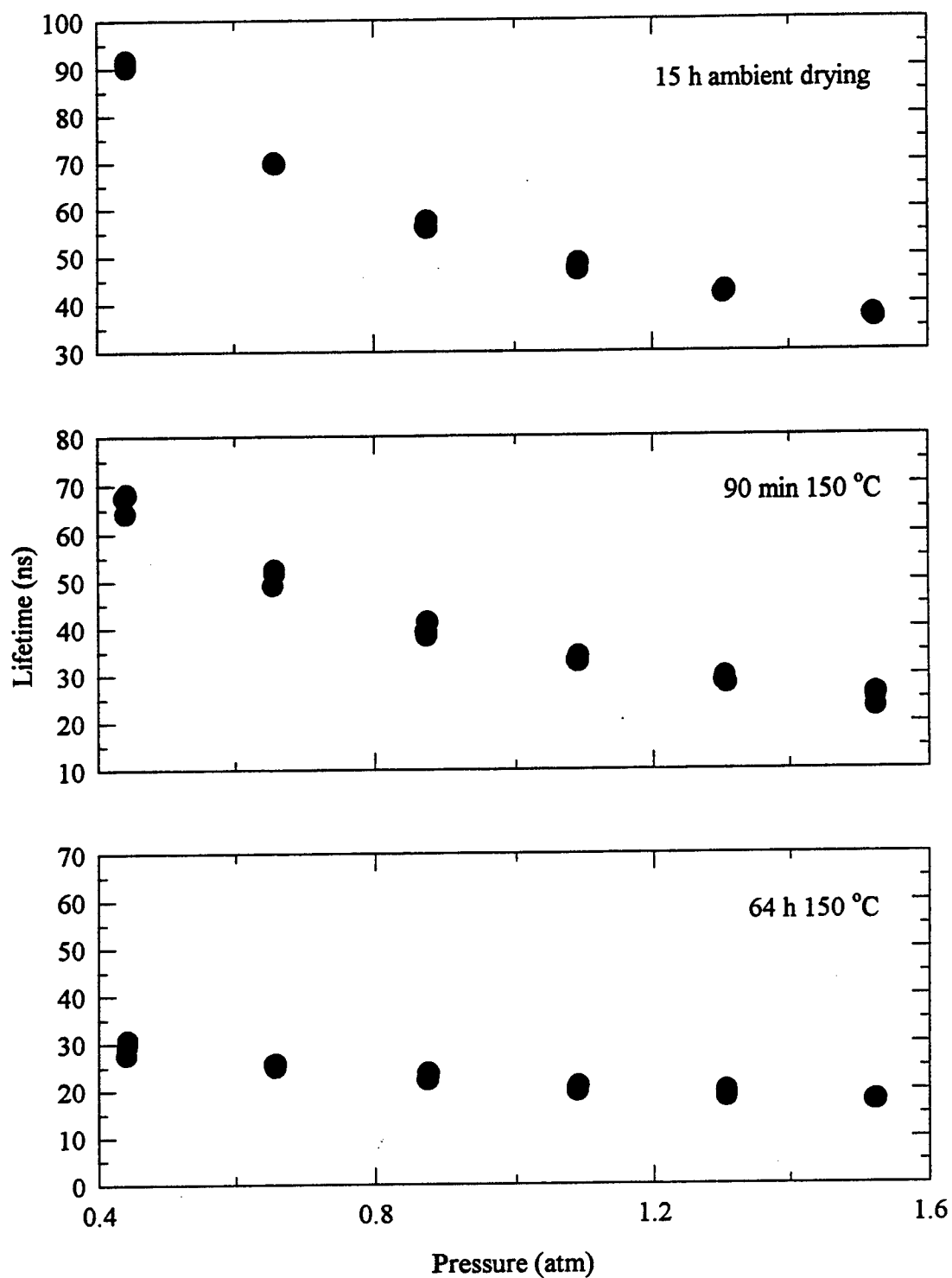
Figure 5. Pressure sensitivity of silicone-based paint as a function of heat treatment (150 ° C).
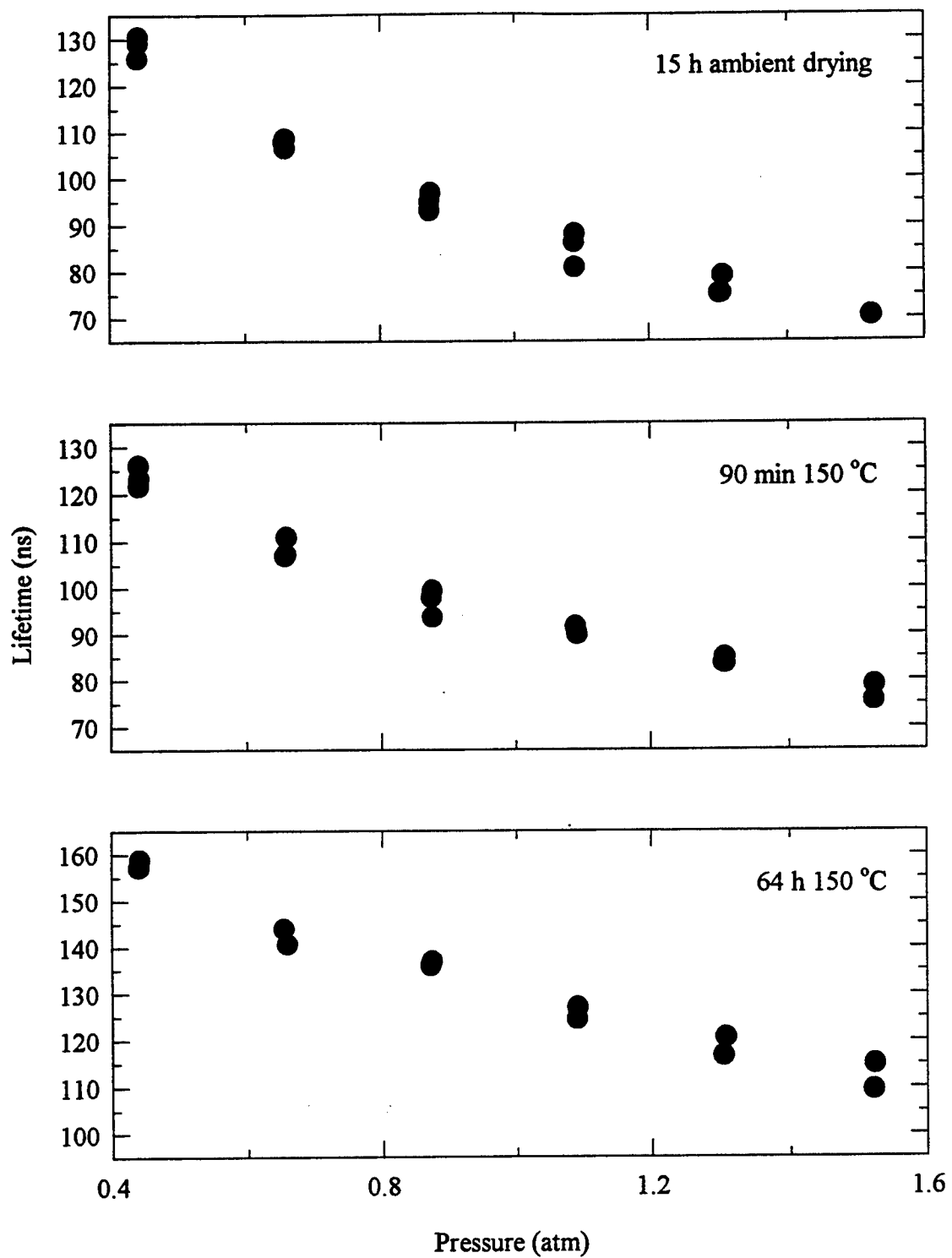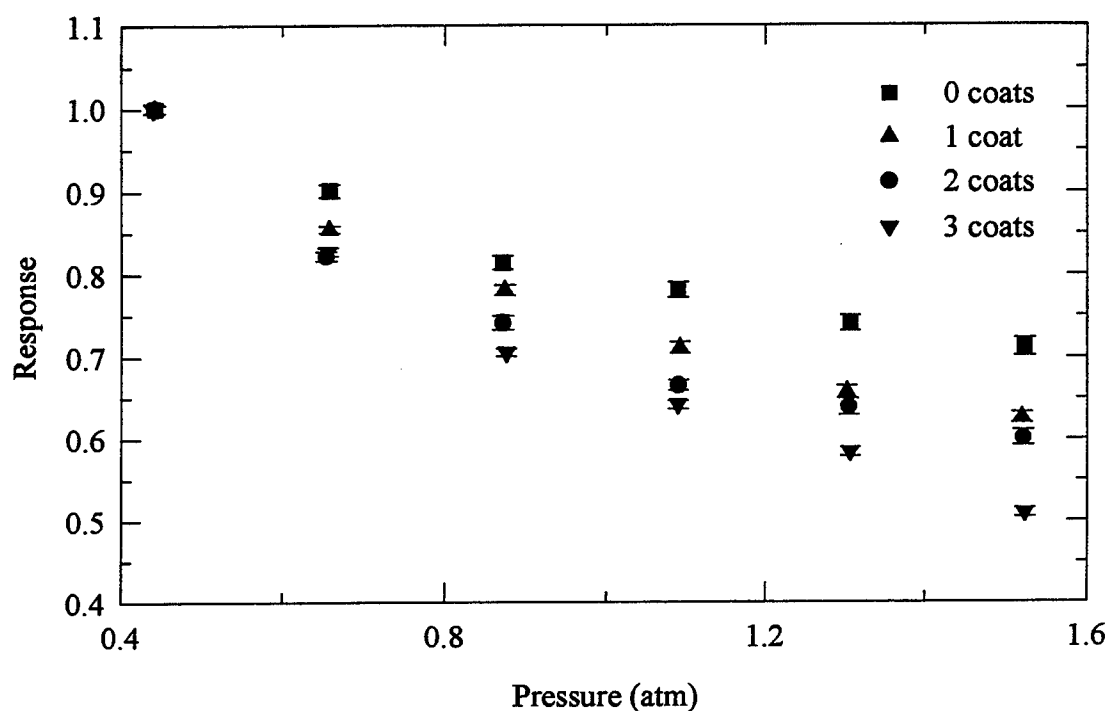
Figure 6. Pressure sensitivity of sol-gel-derived sandwich coating as a function of heat treatment (150 ° C).

The ability to *tune* the pressure sensitivity of a composite coating is particularly attractive. This approach would allow the development of a general thin-film architecture, in which tunable pressure sensitivity was effected by simple changes in the preparation protocol. To date however, pressure-sensitive paints which exhibit tunable performance characteristics have not been demonstrated. The most important attribute of the sol-gel-derived sandwich architecture is the fact that O_2 diffusivity through the upper sol-gel film can be regulated simply by changing the thickness (number of coats applied) and/or porosity (chemistry of the sol-gel process). The feasibility of this approach was determined by preparing sol-gel-derived composites consisting of: 1) a lower sol-gel screen layer, 2) a silicone-based active layer, and 3) sol-gel overcoat layers ranging in number between zero and three. The response curves for the resulting samples are shown in Figure 7.

Figure 7. Pressure sensitivity of sol-gel-derived sandwich films as a function of overcoat thickness.

This preliminary data clearly indicates that the pressure-sensitivity of these sandwich coatings can be easily manipulated by adjusting of the thickness of the sol-gel overcoat layer. This approach, coupled with control over the sol-gel process, provides a convenient and effective means of modulating the composite pressure sensitivity for specific test conditions of interest.

The temporal response of a pressure-sensitive coating will ultimately dictate the range of transient phenomena which can be probed. For many turbomachinery applications, typical transient events occur on a 1 kHz time scale. Thus, their investigation requires a composite coating which exhibits ms or sub-ms response times ($\leq$ 10^{-3} s). In order to determine the temporal response of the sol-gel-derived sandwich films, the following experiment (Figure 8) was designed and implemented.
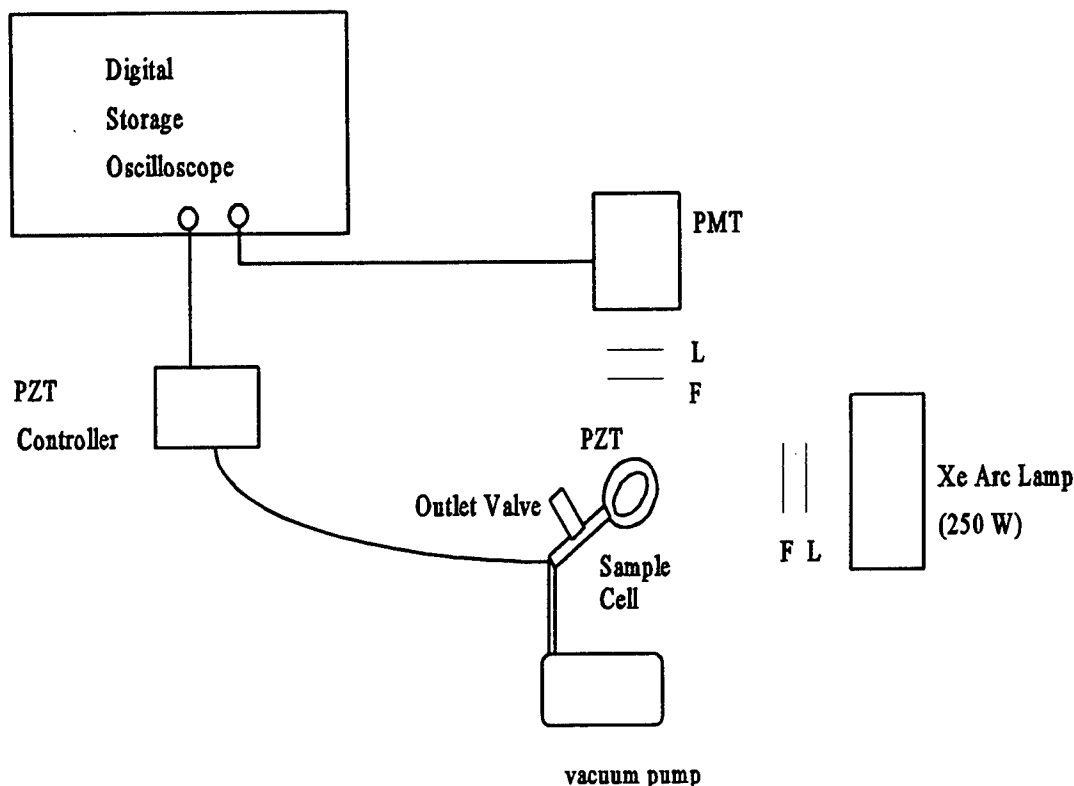
Figure 8. Instrumental apparatus for temporal response determination.

A piezoelectric transducer (PZT) and the sample (1 cm diameter aluminum coupon) were sealed to atmosphere using a circular (1 cm^2) fused silica window and rubber o-ring within a circular brass cell (2.5 cm diameter). The output from a 250 W Xe arc lamp was filtered (337 nm notch filter) and focused onto the pressure-sensitive coating. Fluorescence from the sample was filtered (380 ± 20 nm bandpass filter) and focused onto the face of a fast-wired PMT for detection. In operation, the cell was evacuated using a mechanical pump and the appropriate plumbing components. Tygon tubing placed over the cell outlet allowed the vacuum to be maintained. Release of the Tygon tubing initiated a rapid rise in cell pressure. The signal from the PZT was used to trigger the digital storage oscilloscope, and the PZT and PMT signals were digitally recorded. Typical results are shown in Figure 9.

17-13
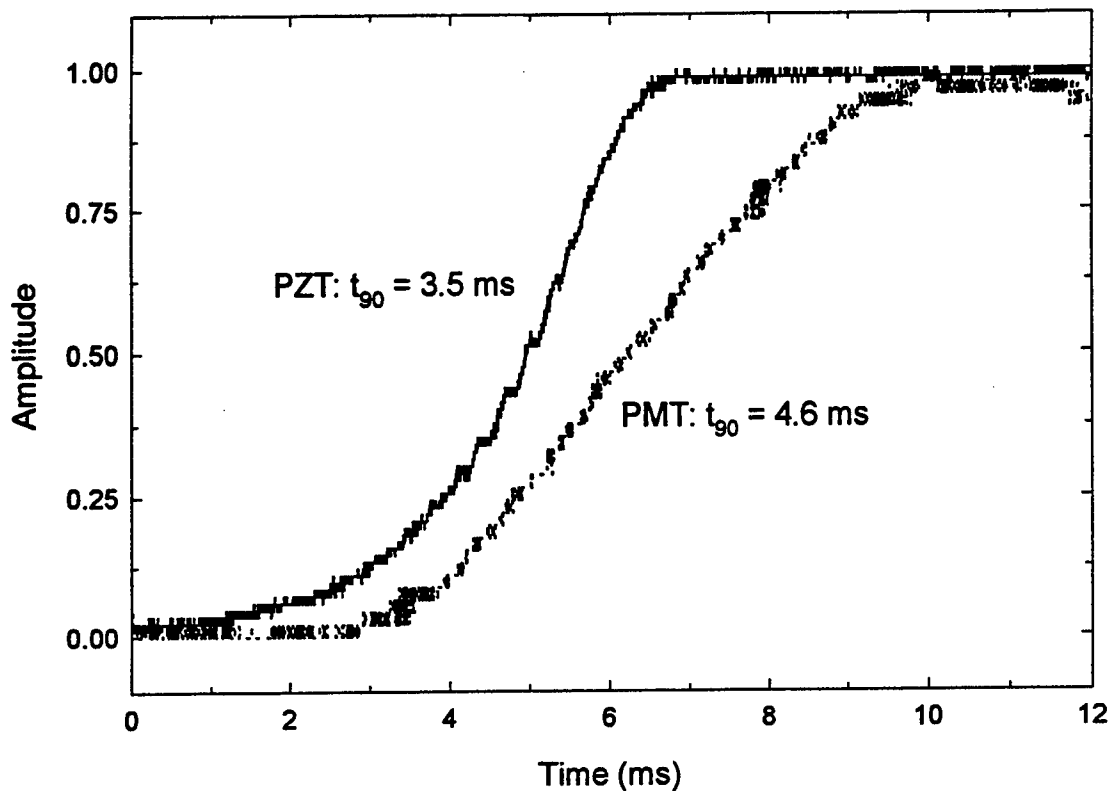
Figure 9. Temporal response behavior of sol-gel-derived sandwich films.

Temporal response values were calculated from this data by determining the time required to reach 90 % of the total response (t_{90}). The results indicate that the temporal response of the sandwich coatings is approximately 1 ms, and is thus amenable for probing transient phenomena which occur on a 1 kHz time scale.

The imaging capabilities of the sol-gel-derived pressure-sensitive coatings were investigated using an air jet impinging on an aluminum surface. The instrumental setup is outlined in Figure 10. In this approach, output (337 nm) from a nitrogen laser was filtered and focused into the proximal end of a 1 mm optical fiber and directed toward the sample. Single-shot images were obtained using an intensified charge-couple detector (ICCD) which was triggered by the laser pulse. A pressure gradient was created across the surface of the sample (2" x 3" aluminum block) using a 10 psi jet of air. Images were collected and stored using a personal computer (PC) and software native to the ICCD camera. The resulting image shown in Figure 10 was generated by dividing the
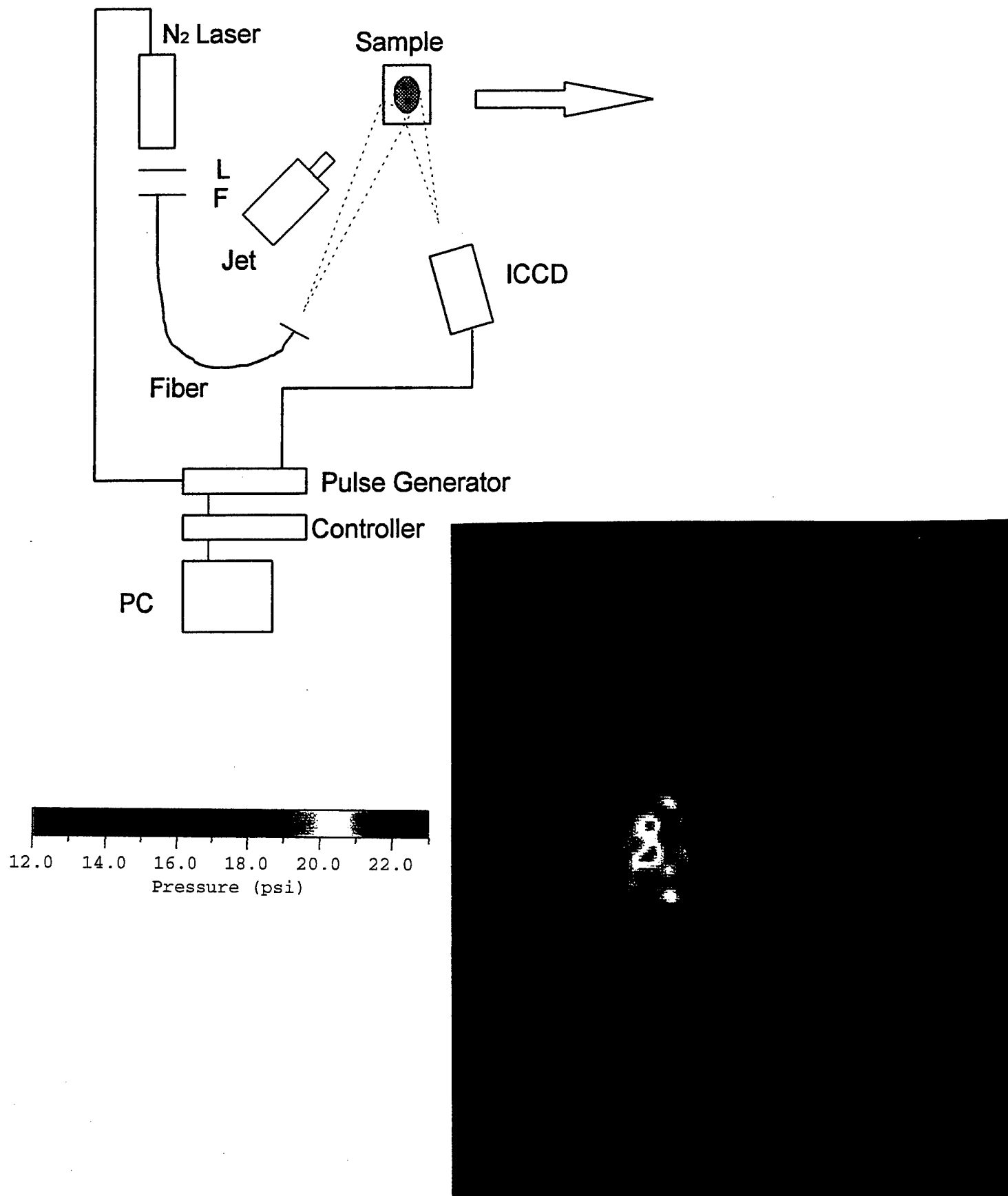
Figure 10. Fluorescence image of a jet (10 psi) impinging on an aluminum surface.

intensity image obtained at ambient pressure ("wind off") by the image obtained under the 10 psi pressure field. The calibration scale provided was obtained by off-line calibration of the pressure-sensitive coating. This image illustrates three points. First, the region in the image where the intensity *ratio* is the highest corresponds to the point of jet impingement, and the diminution in the intensity ratio away from this area reflects the pressure gradient formed at the surface. Second, the dark region on the left edge of the image reflects the sub-ambient pressure field resulting from the jet impinging on the test surface at an angle. Third, the visualization of pressure fields above ambient (1 atm) was not possible with the silicone-based paint alone, indicating that the sol-gel-derived overcoat layer improved the high-pressure sensitivity of the composite coating.

The identical optical sampling method was then used for a fuel-injector application in a hypersonic tunnel. For this work, an elliptical injector nozzle was machined into a 100 cm^2 aluminum block. Liquid flow was introduced from the bottom of the sample block at 95.2 psi, and emerged into an air stream (mach 2) which was orthogonal to the sample surface. The results are shown in Figure 11 and illustrate several interesting features. First, in general, the relatively low intensity reflects the sub-ambient pressures experienced under hypersonic flow. Second, the high intensity ratio region centered in the lower portion of the image represents the injector nozzle, and the higher pressures (~ 2 atm) encountered at the point where the injector stream breaches the mach 2 flow stream. Third, the lower pressure regime emanating from bottom of the image up is the injector plume. Fourth, the large semi-circle spanning the image is a visualization of the oblique shock wave, indicative of a supersonic flow impinging on a surface. Finally, it should be noted that the sub-ambient pressure fields and low temperatures encountered in this test are ideal for use with the silicone-based paint; however, their use would preclude tunnel heating. Thus, the sol-gel-derived sandwich architecture not only imparts greater temperature stability and high-pressure sensitivity, but is simultaneously responsive to low-pressure regimes.
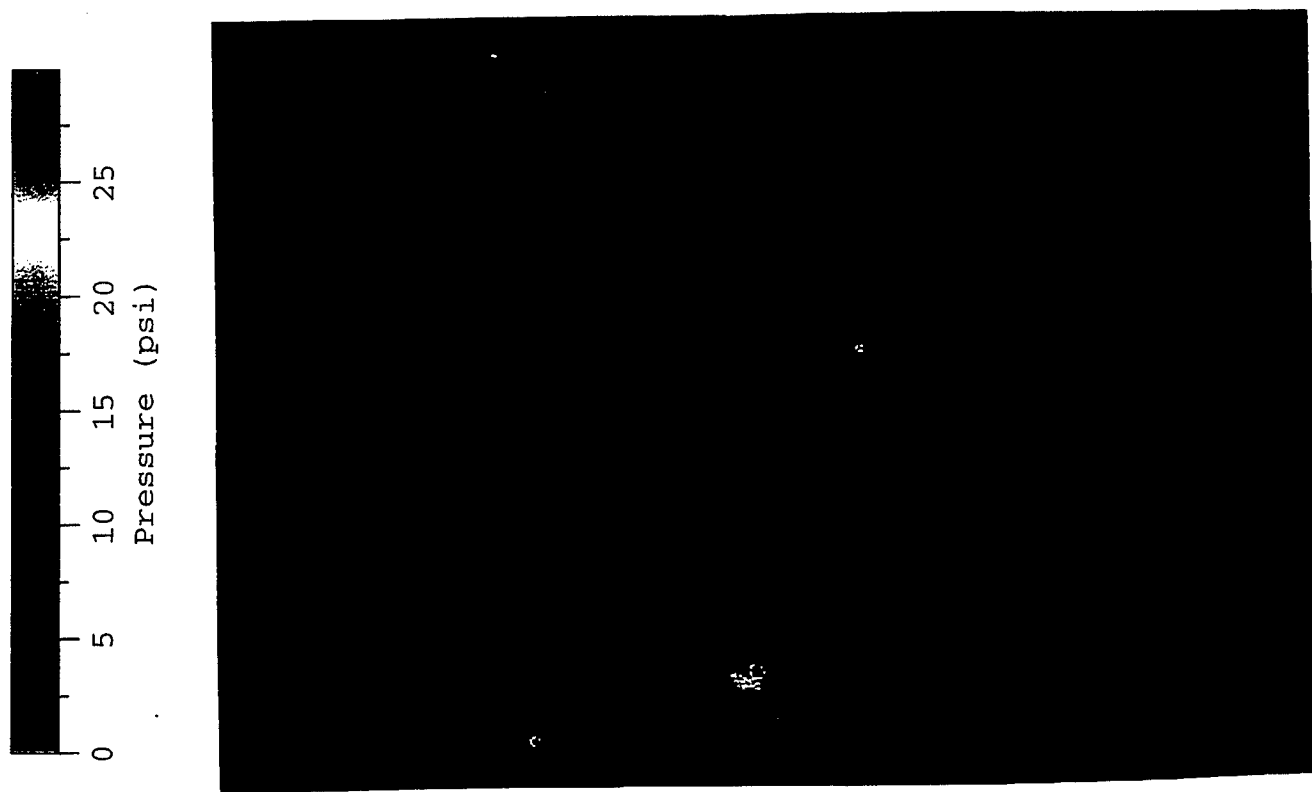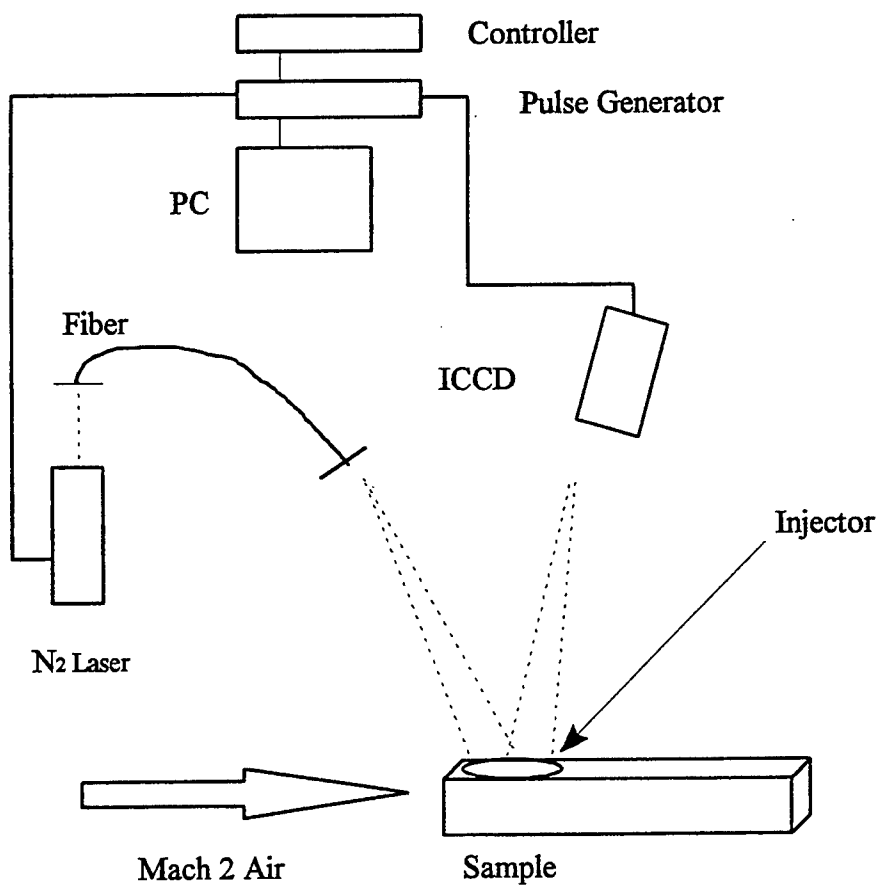
Figure 11. Fluorescence image of an injector (95.2 psi) in a mach 2 wind tunnel.

Conclusions

The generation of spatially-continuous pressure maps of aerodynamic surfaces represents a fundamental requirement for the development of aerospace technology. Current methods involving pressure taps and electronically-scanned transducers are limited to providing pressure information at discrete locations, are expensive to implement, and suffer high-failure rates. The use of pressure-sensitive paint technology represents a viable alternative, but has been impeded by the lack of a general protocol for the production and preparation of coatings exhibiting the desired performance characteristics. The current work demonstrates that sol-gel-derived composites offer a convenient means to produce pressure-sensitive composites that exhibit: 1) temperature stability, 2) *tunable* pressure sensitivity, and 3) ms temporal response. Thus, sol-gel-derived pressure sensitive coatings represent a powerful tool for aerodynamic research that will aid the development and evaluation of CFD models, and benefit both the commercial and military aircraft industries in the United States.

References

1.   J. Kavandi, *et al.*, "Luminescent Barometry in Wind Tunnels," *Rev. Sci. Instrum.* **1990** *61*, 3340.

2.   M. Gouterman, *et al.*, "Surface Pressure Field Mapping Using Luminescent Coatings," *Experim. Fluids* **1993**, *14*, 33.

3.   M. Morris, *et al.*, Aerodynamic Applications of Pressure-Sensitive Paint," Paper No. 92-0264, AIAA 30th Aerospace Sciences Meeting and Exhibit, January 1992.

4.   M. Morris, *et al.*, "Aerodynamic Measurements Based on Photoluminescence," Paper No. 93-0175, AIAA 31st Aerospace Sciences Meeting and Exhibit, January 1993.

5.   J.R. Lakowicz, "Principles of Fluorescence Spectroscopy," Plenum Press: New York, NY, 1983.

6.   J.D. Jordan, *et al.*, "Aerosol-Generated Sol-Gel-Derived Thin Films: Biosensor Platform," *Anal. Chim. Acta,* in press (June 1996).

7.   J.D. Jordan, *et al.*, " Characterization of Aerosol-Generated Sol-Gel-Derived Thin Films," *Chem. Mater.*, submitted for publication (June 1996).

8.   J.D. Jordan, *et al.*, "Development of Sol-Gel-Derived Pressure-Sensitive Coatings for Fluorescence Lifetime-Based Imaging," *Appl. Spectrosc.*, manuscript in progress (September 1996).

9.   U. Narang, *et al.*, "A Novel Protocol to Immobilize Active Urease in a Tetraethoxysilane-Derived Sol-Gel Thin Film Architecture," *Chem. Mater.* **1994**, *6*, 1596.

10.   U. Narang, *et al.*, "Glucose Biosensor Based on a Sol-Gel-Derived Platform," *Anal. Chem.* **1994**, *66*, 3139.

SUPER-CAPACITOR BOOST CIRCUIT
AND
SUPER-CAPACITOR CHARGER

Brett A. Jordan
Graduate Student
Department of Electrical Engineering

Wright State University
Dayton, OH 45435

Final Report for:
Graduate Student Research Program

August 1996

SUPER-CAPACITOR BOOST CIRCUIT
AND
SUPER-CAPACITOR CHARGER

Brett A. Jordan
Graduate Student
Department of Electrical Engineering
Wright State University

Abstract

Large capacitors and their charging circuits have a variety of potential uses in such applications as emergency lighting systems, backup and emergency power supplies, and other applications where battery based systems are not practical. Prior art battery chargers are not suitable candidates for charging these large capacitors because of the high power dissipation and currents encountered at the beginning of the charging cycle.

Our summer research has concluded by our group filing for two separate patents. One patent is based upon a Large Energy-Storage Capacitor Charger. The other patent is based upon a Capacitor-Based Boost Circuit.

Since our group has filed for patents on these circuits, I am prevented from filing a full length report.

A COMPARATIVE STUDY OF NUMERICAL SCHEMES AND TURBULENCE MODELS IN PREDICTING TRANSVERSE JET INTERACTIONS WITH A SUPERSONIC STREAM

Greg Laskowski
Graduate Student
Department of Aerospace Engineering and Engineering Mechanics
University of Cincinnati
Cincinnati, OH 45221-0070

Final Report for:
Summer Graduate Student Research Program

August 1996

A COMPARATIVE STUDY OF NUMERICAL SCHEMES AND TURBULENCE MODELS IN PREDICTING TRANSVERSE JET INTERACTIONS WITH A SUPERSONIC STREAM

Greg Laskowski
Graduate Student
Department of Aerospace Engineering and Engineering Mechanics
University of Cincinnati

Abstract

A numerical simulation investigation was conducted to study the flow field near a normally injected secondary flow into a turbulent, supersonic freestream over a flat plate. Different numerical schemes and turbulence models were assessed in terms of their ability to predict the resulting flow field with its complex shear layers, multiple shocks and subsonic recirculation regions. Two implicit Navier-Stokes solvers, namely, COBALT and NPARC, and four turbulence models, were used in the numerical simulation. The computational results are presented and compared with existing experimental data for the separation point and surface pressure distribution. The results indicate that the best agreement with the experimental data in terms of separation point location were obtained with COBALT using the one equation Spalart-Allmaras turbulence model, while the closest agreement with the separation induced pressure rise upstream of the slot was obtained with NPARC with the two equation k-ω model.

A COMPARATIVE STUDY OF NUMERICAL SCHEMES AND TURBULENCE MODELS IN PREDICTING TRANSVERSE JET INTERACTIONS WITH A SUPERSONIC STREAM

Greg Laskowski

Introduction

New advanced, lighter weight, lower cost nozzle designs are required for future advanced fighter aircraft. One of the approaches considered to achieve the USAF nozzle technology development goals by the year 2000 is to use fluidic control, by injecting secondary air symmetrically for throttling or asymmetrically for thrust vectoring [1]. Numerical simulations can be used to investigate the effects of injection slot location, injection angle, and injected mass flow on the nozzle performance in order to reduce the number of experimentally tested configurations [2]. They can also provide more insight into the detailed nozzle flow characteristics under these conditions. The flow field in the vicinity of the jet injection is very challenging to simulate numerically due to the presence of a Mach surface, multiple shocks, shear layers, flow separation and low speed recirculating flow zones. A high degree of confidence in the predictions of the numerical computations is required, especially for the separation location, since it has a strong influence on the nozzle static thrust coefficient [3]. Furthermore, the onset of flow separation and the extent of the shock induced flow in supersonic flows has been found to be affected by turbulence models and numerical algorithms [4].

In this project, two dimensional numerical computations were conducted to study the flow field resulting from normal jet injection in supersonic flow over a flat plate. The simulations were conducted using two different computer codes for the implicit solution of the compressible Navier Stokes equations and four different turbulence models. The computational results are compared with existing experimental data for the surface pressure distribution and separation location. The one equation Spalart-Allmaras turbulence model in the COBALT code gave separation point location predictions in closest agreement with the experimental data, while the two-equation k-ω model in NPARC gave best overall agreement with the experimental pressure distribution.

Configuration and Flow Conditions

The simulated flowfield is shown schematically in Figure 1[5]. Injecting secondary flow normal to the freestream acts as blockage to the primary flow forming a bow shock also termed the jet induced shock. The static pressure increases due to the induced shock which causes the flow to separate resulting in two counter-rotating recirculating regions and a separation shock. Downstream of the slot, a recirculating zone is formed along with a reattachment shock which redirects the flow parallel to the plate surface. In addition, a characteristic Mach surface is formed where the high pressure jet adjusts to the free stream pressure.

Aso et al. [5] reported experimental results for secondary flow injection normal to the turbulent boundary layer over a flat plate in a supersonic stream. The geometry tested consisted of a flat plate with a sharp leading edge and a slot located 0.330 m downstream of the leading edge. The experimental data consisted of Schlieren photographs and surface pressure distributions obtained at free stream Mach numbers ranging from 3.70-3.81 and unit Reynold's number of 5.83×10^6 /m. Measurements were obtained for slot width's of 0.5 mm, 1.0 mm, and 2.0 mm, and injected flow total pressure ratios relative to free stream of 0.083, 0.180, 0.310, 0.440, and 0.510. Presented herein is a comparison of the numerical results with the experimental data for the case of 3.71 free stream Mach number, 1.0 mm slot width, and 0.310 pressure ratio.

Code Descriptions

Two computer codes, namely NPARC [6] and COBALT [7], were used to obtain the numerical solution to the compressible two dimensional Navier-Stokes equations. NPARC solves the Reynolds averaged Navier-Stokes equations in conservation law form and general curvilinear coordinates using a central difference, second order accurate algorithm based on the approximate factorization scheme of Beam-Warming. The algorithm forms an ADI type of scheme in which each sweep involves the inversion of a set of scalar pentadiagonal matrices. Second and forth order dissipation modeled after Jameson is included to increase the robustness of the code. Version 2.2β of the code, which was used in the present investigation, incorporates six different turbulence models, namely Baldwin-Lomax, Baldwin-Thomas, Baldwin-Barth [8], and RNG one equation algebraic models, and the two equation k-ε and k-ω turbulence models of Chien [9] and Wilcox [10] respectively.

COBALT uses a finite-volume, cell-centered, first-order accurate in space and time, exact Riemann solver of Godunov [11]. The approximate Riemann solution method due to Collela [12] in combination with the iterative method of Gottlieb and Groth [13] is used. Second-order accuracy in space is patterned after van Leer's [14] MUSCL scheme where the flow is assumed to vary linearly within each cell. The linear variations are constructed by a least squares method that is solved by QR factorization. Second-, third- and fourth-order temporal accuracy is achieved via the low-storage methods of Williamson [15]. Second-order accurate viscous terms patterned after McCormick [16] are added to yield a Navier-Stokes solver. The one equation Spalart-Allmaras [17] and Baldwin-Barth turbulence models are available in the code. The code was developed primarily for unstructured grids but is applied to a structured grid in this project.

Computational Details

The computational domain used in assessing the two codes is similar to that used by Rizetta [18]. The upper boundary was placed far enough from the plate surface such that all shocks passed through the exit boundary. The downstream boundary was placed far enough downstream of the reattachment point.

Figure 2 represents the 253x81 grid used in the numerical solutions which was generated using the software package GRIDGEN [19]. Ten equally spaced points were placed ahead of the plate leading edge to accommodate the use of the free-stream conditions at the in-flow boundary. A total of 120 points were distributed from the plate leading edge to the slot. Within the slot, 15 equally space points were placed corresponding to 16 segments of $\Delta x/l = 1.89 \times 10^{-4}$. The grid was clustered near the slot in the streamwise direction using algebraic stretching to ensure adequate resolution near the injection location. Normal to the plate surface the grid was stretched such that $\Delta y/l_{min} = 1.0 \times 10^{-5}$ corresponding to a y^+ value of about 1.0 which is necessary for producing good results with the k-ω model [20] and well below the recommended value of 3.0-5.0 for both the Baldwin-Barth [8] and Spalart-Allmaras models [17].

The free stream static pressure, temperature and Mach number were specified at the in-flow boundary. Symmetry condition was invoked at the lower boundary connecting the incoming flow and the leading edge of the flat plate. No-slip, adiabatic wall condition was specified over the plate surface. In the NPARC solution, first order extrapolation of all flow variables was applied at the upper and exit boundaries whereas for COBALT the Riemann invariant boundary condition was invoked at the same boundaries.

One of the fundamental differences between the two codes was in the specification of the boundary conditions across the injection slot opening. NPARC required specification of the stagnation pressure and temperature at each grid point whereas COBALT required specification of the static pressure, temperature and Mach number across each cell surface. The appropriate ratios between stagnation and static pressure and temperatures at choked conditions were used to specify uniform equivalent flow conditions in each case.

Turbulence quantities for the simulation were not readily available. Therefore, first order extrapolation was applied since it was the most straightforward approach for maintaining consistency between the two codes and different turbulence models.

Starting from uniform initial conditions corresponding to free stream conditions, the solution was advanced using local time stepping. For NPARC, the CFL number was varied from 0.2 initially to 0.6 at latter stages in the computations. For COBALT, the CFL number was determined within the explicit solution throughout the entire iterative process. A total of 30,000-40,000 iterations were necessary for NPARC while 70,000-80,000 iterations were performed using COBALT. The solution was considered converged when the maximum change in the friction coefficient was < 0.5% after 2000 consecutive iterations, which occurred directly downstream of the separation point. In the case of NPARC, this corresponded to 3 to 4 orders of magnitude reduction in the total L2 residual.

Results and Discussions

Figure 3 presents the Mach number contours from the solutions using COBALT and NPARC omitting the Baldwin-Thomas model due to its poor prediction of separation and surface pressure. The

figure clearly shows the jet induced and separation shock upstream, and the reattachment shock downstream, which is especially well defined in the NPARC k-ω case. The general shape of the plume is in good agreement between the two codes and different turbulence models.

Figures 4 and 5 compare the computed surface pressure distributions and surface friction using the two codes and four turbulence models to the experimental data. Rizzetta's computational results [18] which were obtained using McCormack explicit scheme and k-ε turbulence model are reproduced in the two figures. From Fig. 4, one can see that the predictions using NPARC with the k-ω turbulence model is closest to the experimental data in terms of the onset of the surface pressure rise associated with the flow separation, and pressure gradient upstream of the injection. COBALT with the Spalart-Allmaras model gave the second best predictions with the pressure rising earlier and at a higher initial rate than the experimental results. The Baldwin-Barth of NPARC predicted the onset of the pressure rise much earlier than the experimental data, however, it most accurately predicted, the syatic pressure level near the slot. COBALT predictor with using the Baldwin-Barth model are close to Rizzetta's with the onset of the pressure rise shortly behind the experimental data, and a much higher shock strength than the experimental data and all other predictions except for NPARC with the algebraic Baldwin-Thomas model. All computations underpredicted the pressure recovery behind the slot, but COBALT predictions were the closest to the experimental results further downstream.

The computed skin friction coefficients, are presented in Figure 5, and the predicted separation locations are compared to the experimental results in Table 1. According to this data, the computed separation location using COBALT with the Spalart-Allmaras model is in closest agreement with experimental results (within 0.14%) followed by NPARC with the k-ω model (within 0.53%). The Baldwin-Barth model with NPARC and COBALT resulted in overpredicting and underpredicting the separation location by 1.84% and 1.81% respectively, with Rizzetta's results very close to the latter. Again, the Baldwin-Thomas algebraic model resulted in the greatest error in predicting the separation location, deviating by more than 3%. Further investigation of the effect of grid refinement study may be required to resolve the spread in the computed friction coefficients ahead of the separation region.

Immediately downstream of the slot there is a noticable difference between the friction coefficient distributions predicted by COBALT and NPARC. Comparing the velocity vectors in the recirculation zone downstream of the slot for the two codes and Baldwin-Barth model as compared in Figures 6 and 7, one can see only one recirculation in the COBALT predictions, but while NPARC predicts a secondary smaller counter-rotating recirculation zone can be seen in NPARC predictions. It was determined in the present investigation that this second vorticy was formed irrespective of the turbulence model, pressure ratio, slot width, artificial viscosity, and CFL number. Yeneriz et al. reported similar findings concerning this type of vortical structure in their PARC code predictions of hypersonic flow over an axisymmetric cone with injection [22]. However, Rizzetta also reported only one recirculation region downstream of the slot as in the COBALT predictions.

The computed static pressure, density, Mach number and momentum distribution across the jet are presented in Figures 8-11. The mass and momentum flux were numerically integrated across the slot opening and the results are presented in Table 1. The flow and thrust coefficents, which represent the ratio between the computed and ideal sonic jet, were found to be in agreement withing 99.99% and 98.31% respecively. According to the computational results COBALT predicts higher pressures and densities and lower Mach numbers and momentum compared to NPARC. The difference between the predictions of the two codes is associated with how the boundary conditions at the slot were implemented in each code.

	COBALT Baldwin-Barth	COBALT Spalart-Allmaras	NPARC Baldwin-Barth	NPARC Baldwin-Thomas	NPARC k-ω	Experiment/Ideal
Separation Point (x/l)	0.917	0.899	0.884	0.929	0.905	0.900
Percent Error	*+1.84*	*+0.14*	*-1.81*	*-3.11*	*+0.53*	-
Jet Mass Flux (kg/s/m)	0.8989	0.8990	0.8994	0.8993	0.8994	1.0278
Flow Coefficient	*0.8750*	*0.8750*	*0.8751*	*0.8750*	*0.8751*	-
Jet Momentum Flux (J)	*236.4804*	*236.5067*	*240.6067*	*240.5800*	*240.6067*	*275.4461*
Thrust Coefficient	*0.8585*	*0.8586*	*0.8735*	*0.8734*	*0.8735*	-

Table 1. Comparison of different turbulence models with experimental results.

Conclusions

Two dimensional numerical simulations were conducted to study the flow field resulting from normal jet injection in supersonic flow over a flat plate. The simulations were conducted using two different computer codes for the implicit solution of the compressible Navier-Stokes equations and four different turbulence models. The computational results were compared with existing experimental data for the surface pressure distribution and separation point location. The results indicated that the one equation Spalart-Allmaras model which was implemented in the COBALT code gave the closest agreement with the experimental data in terms of separation point location. On the other hand, the k-ω model which was implemented in the NPARC code gave the best overall agreement with the experimentally measured surface pressure rise upstream of the slot. All computational resulted underpredicted pressure recovery behind the slot.

Acknowledgments

References

1. Gridley, M. C., and Walker, S. H., "Propulsion Integration Issues for Advanced Fighter Aircraft," Air-Force Technical Report 142-96, Wright Laboratory, Wright-Patterson AFB, OH.

2. Wing, D. J., "Static Investigation of Two Fluidic Thrust-Vectoring Concepts on a Two-Dimensional Convergent-Divergent Nozzle," NASA Technical Memorandum 4574, December 1994.

3. Hamed, A., Vogiatzis, C., and Yeuan, J. J., "Performance Prediction of Overexpanded 2DCD Nozzles," 12th International Symposium of Air Breathing Engines, September 1995.

4. Hamed, A., "An Investigation of Oblique Shock/Boundary Layer Interaction Control," AFOSR Technical Report Contract No. 91-0101, January 1992.

5. Aso, S., Okuyama, S., "Experimental Study on Mixing Phenomena in Supersonic Flows with Slot Injection," AIAA Paper No. 91-0016, January 1991.

6. Cooper, G. K., and Sirbough, J. R., "PARC Code: Theory and Usage," Arnold Engineering Development Center Report AEDC-TR-89-15, Arnold AFS, December 1989.

7. Strang, W. Z., "COBALT User's Manual," WL/FIMC WPAFB, OH, February 1996.

8. Baldwin B. S., and Barth T. J., "A One-Equation Turbulence Model for High Reynold Number Wall-Bounded Flows," AIAA Paper No. 91-0610, January 1991.

9. Chien, K.-Y., "Prediction of Channel and Boundary-Layer Flows with a Low Reynolds Number Turbulence Model," AIAA Journal, Vol 20, pp 33-38, January 1982.

10. Wilcox, D. D., "The Remarkable Ability of Turbulence Model Equations to Describe Transition," 5th Symposium on Numerical and Physical aspect of Aerodynamic Flow, Long Beach, CA, January 1992.

11. Gudunov, S. K., "A Finite Difference Method for the Numerical Calculation of Discontinuous Solutions of the Equations of Fluid Dynamics," Mat, Sb., Vol 47, pp 271-290, 1959.

12. Collela, P., "Glimm's Method for Gas Dynamics," SIAM Journal of Scientific Statistical Computing, Vol 3, No. 1, pp 76-110, 1982.

13. Gottlieb, J. J., and Groth, C. P. T., "Assessment of Riemann Solvers for Unsteady One-Dimensional Inviscid Flows of Perfect Gasses," Journal of Computational Physics, Vol 78, pp 437-458, 1988.

14. van Leer, B., "Towards the Ultimate Conservative Difference Scheme. V. A Second-Order Sequel to Godunov's Method," Journal of Computational Physics, Vol 32, pp 101-136, 1979.

15. Williamson, J. H., "Low-Storage Runge-Kutta Schemes," Journal of Computational Physics, Vol 35, pp 48-56.

16. MacCormack, R. W., "The Effect of Viscosity in Hypervelocity Impact Cratering," AIAA Paper No. 69-0354, 1969.

17. Spalart, P. R., and Allmaras S. R., "A One-Equation Turbulence Model for Aerodynamic Flows," AIAA Paper No. 92-0439, January 1992.

18. Rizzetta, D. P., "Numerical Simulation of Slot Injection into a Turbulent Supersonic Stream," AIAA Journal, Vol 30, No. 10, October 1992.

19. Steinbrenner, J. P., Chawner, J. R., and Fouts, C. L., "The GRIDGEN 3D Multiple Block Grid Generation System," WRDC-TR-90-3022 Vols. I and II, Wright-Patterson AFB, OH, 1990.

20. Hamed, A., and Vogiatzis, C., "Assessment of Turbulence Models in Overexpanded 2D-CD Nozzle Flow Simulations," AIAA Paper No. 95-2615, 1995.

21. Yeneriz, M. A., Davis, J. C., Cooper, G. K., and Harvey, D. W., "Comparison of Calculation and Experiment for a Lateral Jet From a Hypersonic Biconic Vehicle," AIAA Paper No. 89-2548, July 1989.
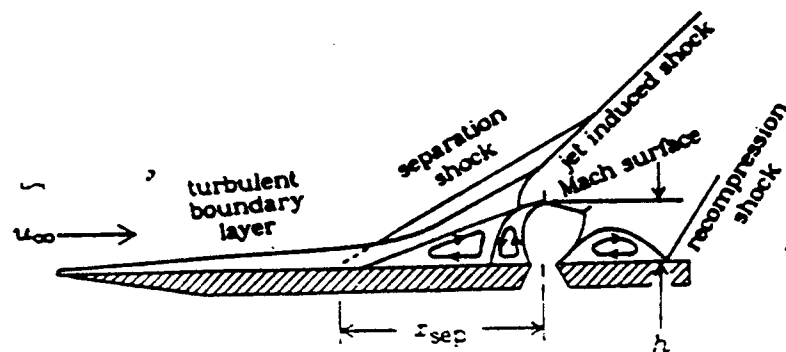
Figure 1. Schematic of flowfield interaction characteristics.
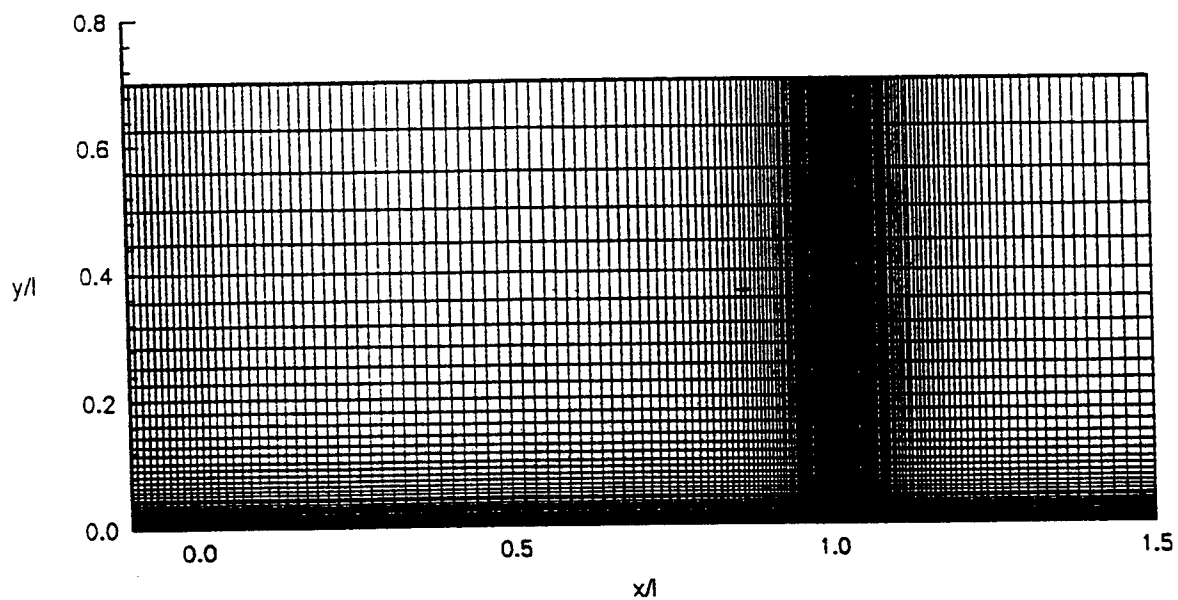
Figure 2. Computation grid for entire flow domain.
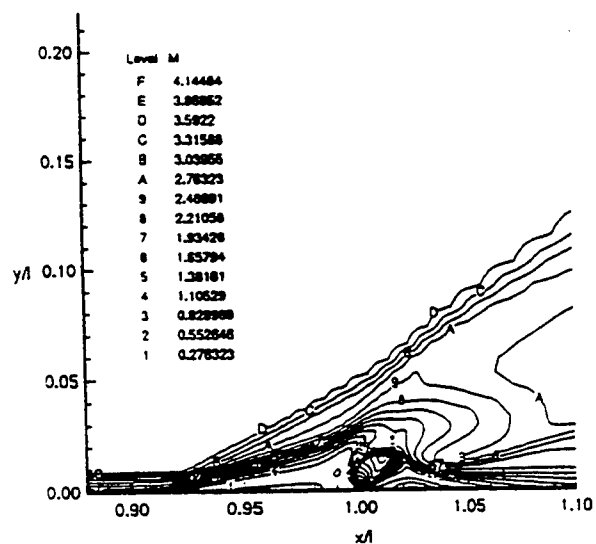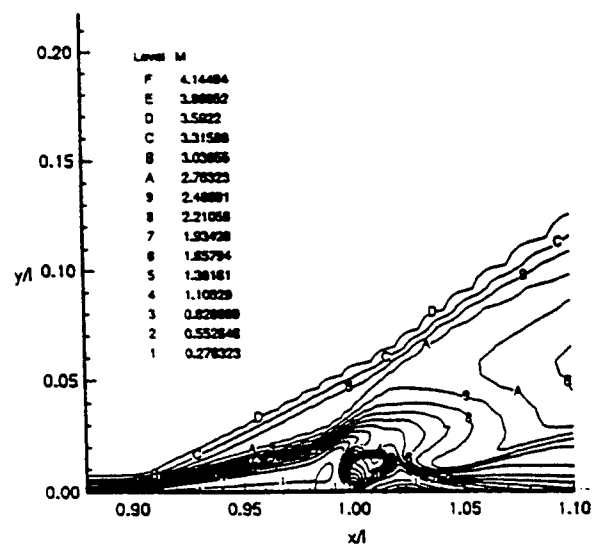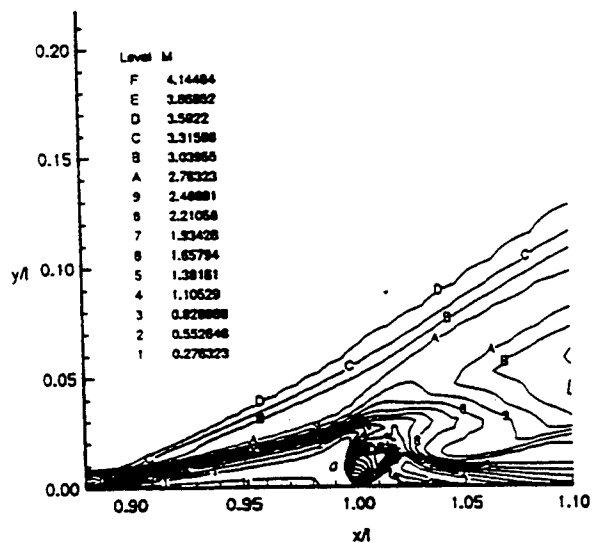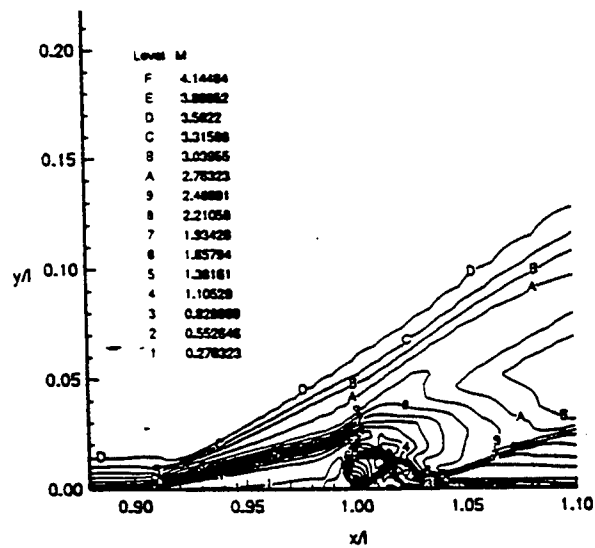
Figure 3. Mach number contours near slot opening. (a) COBALT, Baldwin-Barth (b) COBALT, Spalart-Allmaras (c) NPARC, Baldwin-Barth (d) NPARC, k-ω.
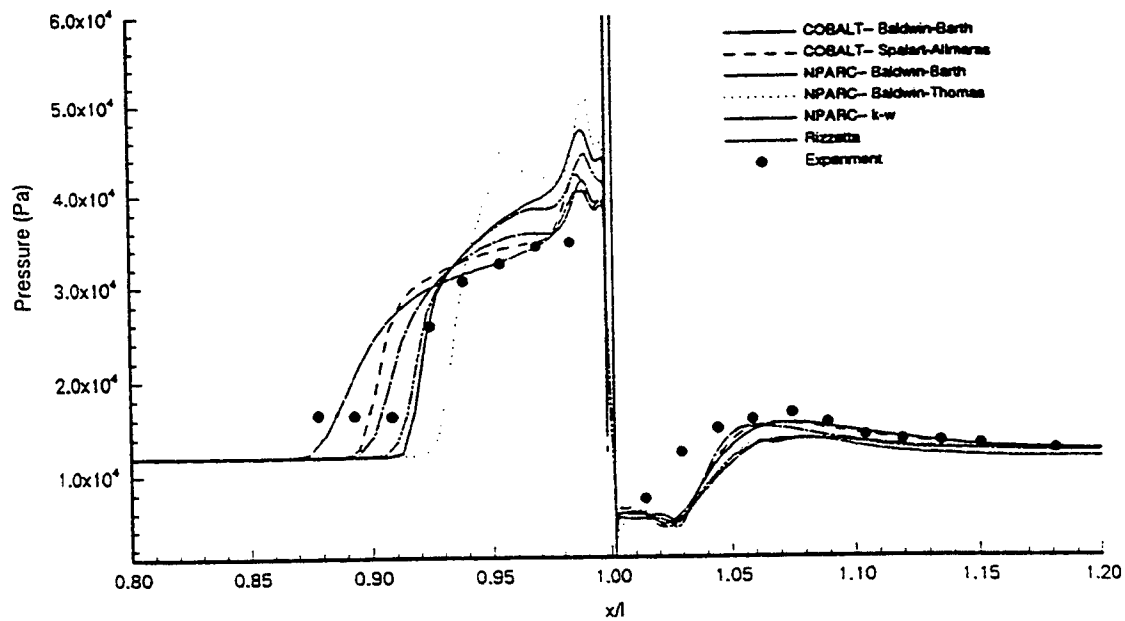
Figure 4. Surface pressure distribution near slot opening.
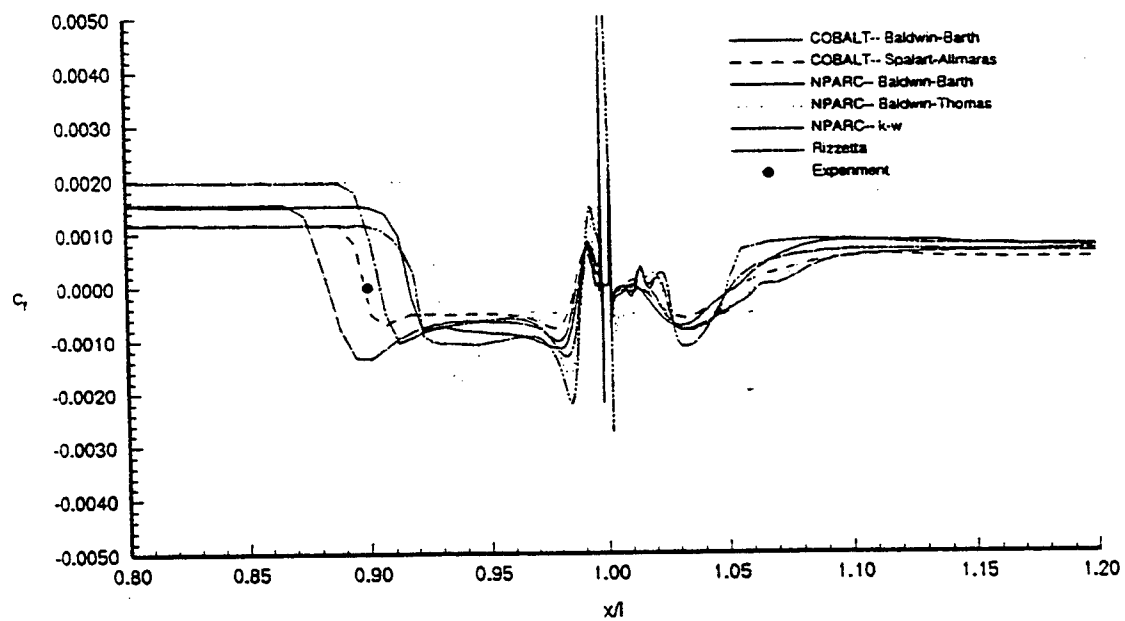
Figure 5. Friction coefficient near slot opening.
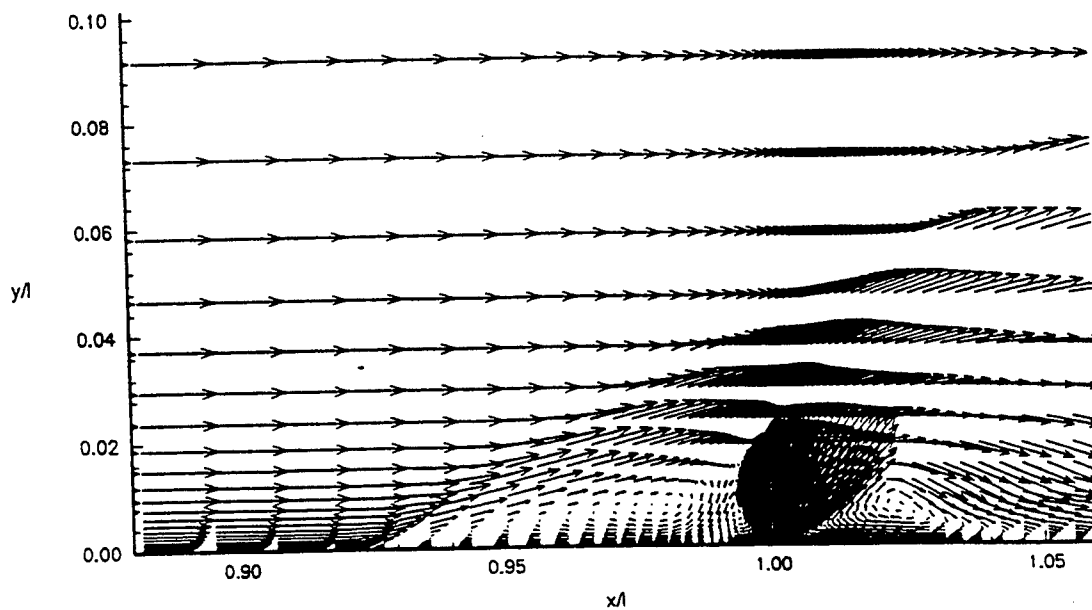
Figure 6. Velocity vectors near slot opening for COBALT using the Baldwin-Barth turbulence model.
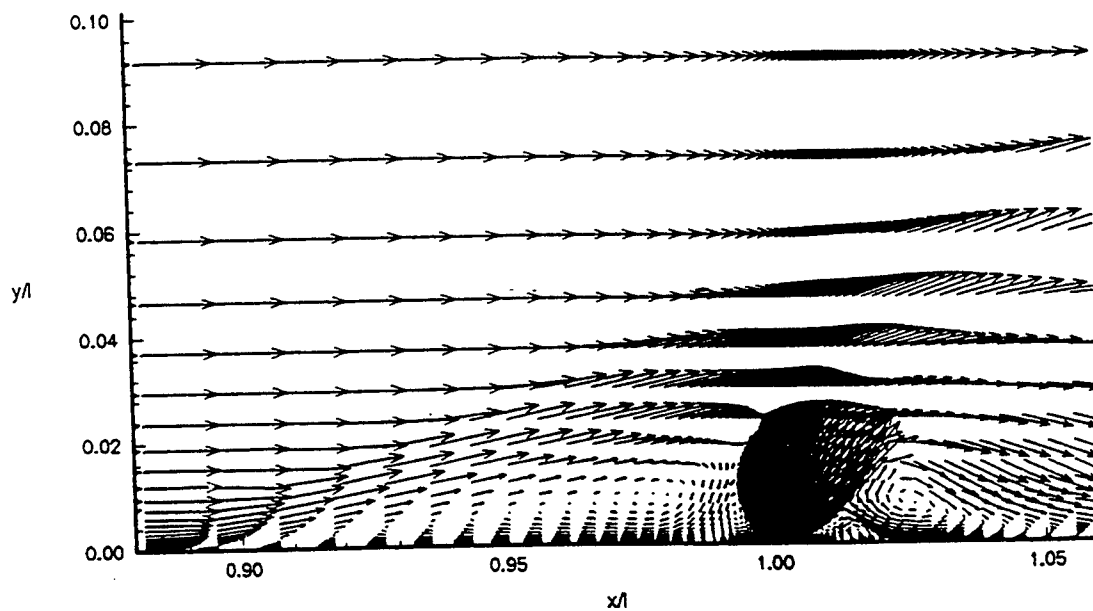
Figure 7. Velocity vectors near slot opening for NPARC using the Baldwin-Barth turbulence model.
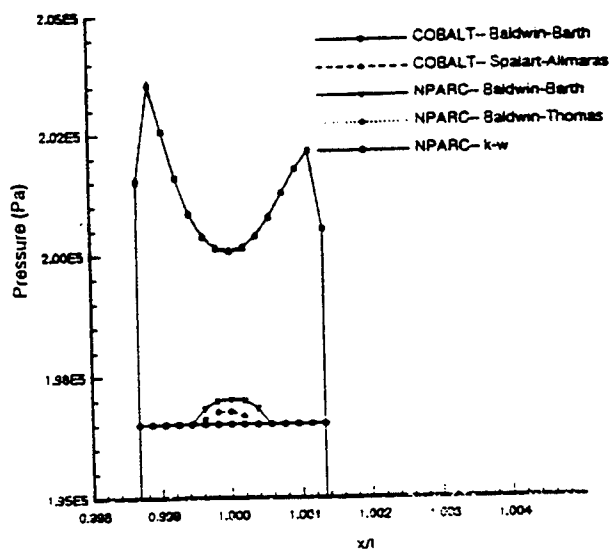
Figure 8. Pressure distribution across slot opening.
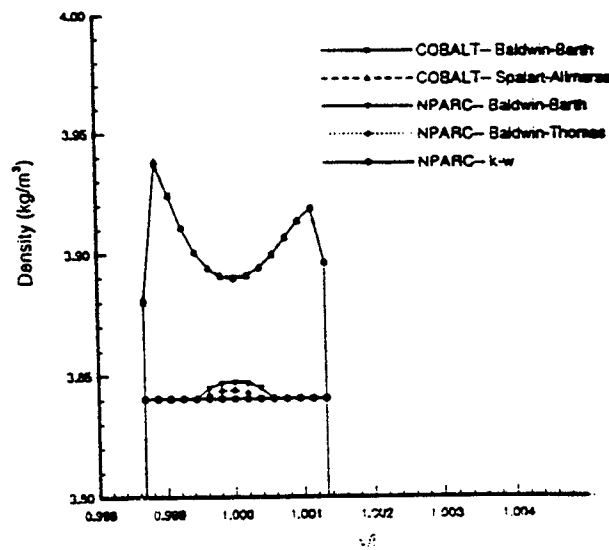
Figure 9. Density distribution across slot opening.
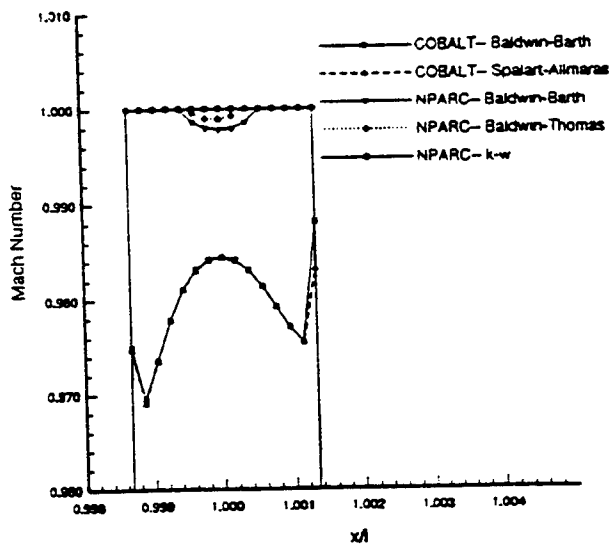
Figure 10. Mach number distribution across slot opening.
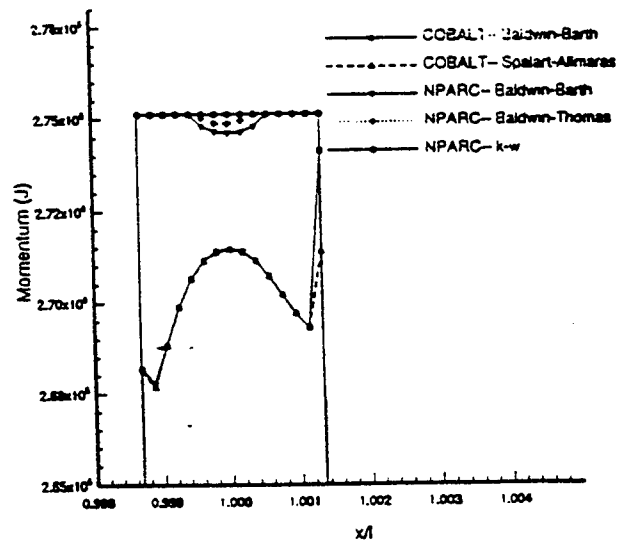
Figure 11. Momentum variation across slot opening.

EFFECT OF HEAT TREATMENT
ON CYCLIC BEHAVIOR OF Ti-22Al-23Nb

Stephanie Luetjering
Graduate Student
Department of Materials Engineering

University of Dayton
300 College Park
Dayton, OH 45469-0240

EFFECT OF HEAT TREATMENT
ON CYCLIC BEHAVIOR OF Ti-22Al-23Nb

Stephanie Luetjering
Graduate Student
Department of Materials Engineering
University of Dayton

Abstract

This research is part of a study in which the main objective is to get a comprehensive understanding of the dependence of fatigue properties on microstructure for the orthorhombic-based titanium aluminide alloy Ti-22Al-23Nb (at.%). Cyclic properties such as the total fatigue life (up to 10^7 cycles), nucleation life, microcrack propagation and macrocrack growth life, will be determined at 23°C and 540°C. In addition, the effect of environment is to be established by performing tests under both air and vacuum conditions at elevated temperature.

In this initial state of the project, the material has been characterized by determining the beta transus temperature and by selecting the heat treatment parameters for two desired microstructures. These microstructures differ mainly in the content of the ordered phases (α_2, beta, and orthorhombic) present. Thus far tensile testing at both temperatures and fatigue testing at room temperature, the latter to establish an SN-curve, has been accomplished for only one of the microstructures. Fractographic analysis has been used to correlate test results with the corresponding microstructural features.

EFFECT OF HEAT TREATMENT
ON CYCLIC BEHAVIOR OF Ti-22Al-23Nb

Stephanie Luetjering

Introduction

The need for low density structural materials with high temperature strength and low temperature ductility for advanced aerospace applications has driven great interest in titanium aluminides, especially in those based on the Ti-Al-Nb system. Titanium aluminide alloys containing the ordered orthorhombic (O) phase, based on the Ti_2AlNb composition /1/, have been found very promising as potential material for the aerospace industry, both in monolithic form and as matrix material for metal matrix composites (MMCs). The benefits of "O" alloys are higher specific strength at elevated temperatures, improved creep resistance, improved fracture resistance and reduced fiber/matrix reaction than state of the art Ti_3Al based alloys, /2 - 7/.

A strong dependency of mechanical properties on microstructural features such as phase constituency, grain size and morphology was determined for various alloy compositions based on the Ti_2AlNb compound, /2, 3, 8/. Primary interest in improving the creep behavior of the investigated alloys focused recent research efforts in this area. The cyclic properties such as fatigue life and crack growth have not been studied in any detail with respect to the material's microstructure. The main objective of this study therefore is to investigate the effect of microstructural changes on the cyclic behavior of a representative ordered orthorhombic-based titanium aluminide alloy, nominally Ti-22Al-23Nb (at.%).

Experimental Procedure

The starting material in this study was sheet-material with a thickness of 0.5 mm of the orthorhombic titanium alloy Ti-22Al-23Nb (at.%). The sheet had been processed from a plasma hearth melted ingot by forging above the beta transus, followed by subtransus pack cross rolling.

Dog-bone specimens were wire electrodischarge machined (EDM) according to Figure 1. The two microstructures were developed by heat treating the specimens in a vacuum furnace (10^{-6} torr) at 20°C below the beta-transus temperature (microstructure A) and 20°C above the beta-transus temperature (microstructure B), then direct cooling to 815°C where the specimens were aged. The beta-transus temperature of the alloy was determined by the disappearing phase technique by solutionizing samples at temperature for 2 hours followed by an oil quench. Furthermore, to optimize the microstructures, a cooling rate study was performed, wherein cooling rates from solutionizing temperature to the aging temperature were varied between 28°C and 100°C. The two different heat treatments are as follows:

    (A)   1055°C for 2 hr \ cool @ 50 °C/min \ 815°C for 8 hr \ furnace cool

    (B)   1095°C for 1 hr \ cool @ 50 °C/min \ 815°C for 8 hr \ furnace cool

Before the testing of the specimens, they are first mechanically polished (600 grid) and then electropolished, to get rid off any defects in the specimens surface (e.g. scratches and rolling lines from the sheet processing) and to round off the edges, an attempt to reduce the stress intensity at these sites.

A servohydrolic test system featuring a horizontal design was used for tensile and fatigue testing. The system is described in detail by Hartman and Buchanan /9/ and Hartman and Russ /10/. Temperature control for the elevated temperature tests was accomplished with thermocouples spot-welded onto a dummy specimen which was placed next to the actual test specimen.

Tensile testing was performed at room temperature and 540°C in laboratory air with a constant cross head speed of 0.0084 mm/s. Isothermal fatigue tests were conducted at room temperature in load control with a sine waveform at a stress ratio of 0.1 and a frequency of 10 Hz.

Materials characterization included bulk chemical analysis and microstructural analysis. Metallographic samples, etched with Kroll's reagent, were studied using a Leica 360 FE scanning electron microscope (SEM). Quantitative phase analysis was accomplished by applying an NIH image analysis program to digitally stored backscattered electron SEM micrographs at 750x and 2000x magnification.

Fracture surfaces of the fatigue specimens were examined using the SEM, crack initiation sites were determined and correlated to microstructural features.

Results and Discussion

Chemistry

Bulk chemistry indicates that the actual composition is closer to Ti-21Al-22Nb (at.%) and that the interstitial contents are at an acceptable level, Table 1.

Ti [at.%]	Al [at.%]	Nb [at.%]	N [wt.%]	O [wt.%]
bal	20.8	22.2	0.010	0.087

Table 1: Chemical analysis of bulk material, Ti-22Al-23Nb

Beta-transus Determination

In Figures 2a and b photos of microstructures are shown, resulting from a 2 hour solutionizing treatment at 1060°C and 1080°C respectively with a subsequent oil quench. The microstructure in Figure 2a still contains the black, globular primary α_2 phase, whereas in the microstructure heat treated at the higher temperature α_2 has been totally dissolved, indicating that the beta-transus temperature has been exceeded in the latter case. To further minimize the interval in which the beta-transus temperature lies, samples were heat treated at 1065°C, 1070°C and 1075°C for 2 hours followed by an aging treatment at 815°C. The beta-transus temperature has not been reached at 1070°C, since the α_2 phase is still present in the corresponding microstructure. After a solutionizing treatment at 1075°C, the α_2 phase only exists near the sheet's surface, as can be seen in Figure 3. This can be explained with a higher oxygen content in this surface area, probably due to the sheet's processing. Other researchers have shown, that the beta-transus temperature is strongly dependent on the interstitial content: with increasing oxygen content the α_2-phase is stabilized and the beta transus is raised to higher temperatures /11/. According to these results the beta-transus temperature lies at about 1075°C.

Cooling rate

The cooling rate from solutionizing temperature to aging temperature influences the growth of the orthorhombic platelets. The higher the cooling rate, the smaller the size of the orthorhombic platelets. In Figure 3 a cooling rate of 28°C/min results in larger orthorhombic platelets at preferred nucleation sites, which are grain boundaries and α_2 particles. Increasing the cooling rate to 50°C/min prevents these orthorhombic platelets to grow to that extent, Figure 4.

Microstructure

A solutionizing treatment below the beta transus results in a microstructure (microstructure A) containing three ordered phases: α_2, beta (β_0), and orthorhombic (O), Figure 4. Approximately 5 vol.% of the structure is represented by the dark, globular primary α_2-phase. The two phase mixture is comprised of 70 vol.% fine, orthorhombic platelets within a continuous, white β_0 matrix. The grains are elongated in the plane of the cross-rolled sheet, with an average diameter of 150 µm and 75 µm in the short transverse direction. This grain size is same as in the as-received microstructure of the sheet-material.

When the solutionizing temperature lies above the beta transus, a fully transformed structure is exhibited (microstructure B), Figure 5. It consists of a two phase mixture of 75 vol.% O phase and a β_0 matrix. A continuous O phase exists along the prior beta grain boundaries. Orthorhombic platelets at these grain boundaries are large compared to those within the former beta grains, due to heterogeneous nucleation. The O platelets within the β_0 grains are approximately the same size as in microstructure A. The grain size is about 300 µm in diameter and thus larger than the one in microstructure A, since grain growth is not being prevented by α_2-particles.

Tensile Properties

Tensile tests have only been conducted on microstructure A. The material's properties for room temperature (RT) testing and for testing at 540°C, average values of three tests, are listed in Table 2.

20-6

Temperature	Yield Strength [MPa]	UTS [MPa]	Elongation [%]	Modulus [GPa]
RT	1003	1156	6.9	117
540°C	730	911	6.6	100

Table 2: Tensile properties of Ti-22Al-23Nb, microstructure A

Surface cracking is observed at 540°C throughout the whole gage section. Cracking occurs predominantly perpendicular to the stress axis along grain boundaries, but also intragranularly in preferred oriented grains, Figure 6.

Fatigue Properties

The room temperature fatigue results for microstructure A are presented in the Stress-Number of Cycles to Failure (SN) diagram in Figure 7. Further fatigue tests have to be performed at maximum stress levels of 750 and 800 MPa to complete this SN-curve. The test run at a maximum stress of 725 MPa, which equals a fatigue ratio of 72 %, was a run-out (10^7 cycles). The low fatigue life of the specimen tested at 775 MPa can be explained with an incorrect electropolishing process. The specimen must have been polished for a too long time period, since the edges have become very pointed, Figure 8. This causes a reduction in cross-sectional area of about 5 % of this thin specimen, which was not accounted for when running the test. The actual applied maximum stress lies around 820 MPa. Usually the electropolishing process results in a specimen shape like in Figure 9, reducing the cross-sectional area by less than 1 %.

Crack initiation sites are found at the specimen's surface, at either grain boundaries or α_2 particles. Crack propagation takes place along the grain boundaries, going around α_2 particles that are lying within the boundaries.

Summary

The sheet-material of the composition Ti-22Al-23Nb (at.%) was used in this study to investigate the effect of microstructure on its cyclic behavior. The beta-transus temperature of the titanium alloy was determined to be at 1075 $\pm$ 4°C. Two microstructures were selected, a three phase microstructure (microstructure A) consisting of the ordered phases α_2, beta and orthorhombic, and a fully transformed structure (microstructure B), containing only orthorhombic platelets in a beta matrix. High temperature testing showed the effect of environment on the tensile properties of the three phase microstructure. Fatigue tests were conducted for the microstructural condition A at room temperature, revealing, that fatigue cracks initiate at areas of high stress intensity, at grain boundaries or α_2 particles at the specimen's surface and that crack propagation is intergranular. Tensile and fatigue testing for microstructural condition B are currently in progress.

References

/1/   *D. Banerjee, A. K. Gogia, T. K. Nandi and V. A. Joshi:* "A new ordered orthorhombic phase in Ti3Al-Nb alloy", Acta. Met. , vol. 36, pp. 871-882, (1988)

/2/   *R. G. Rowe:* " Recent developments in TiAlNb titanium aluminide alloys ", High Temperature Aluminides and Intermetallics, eds. S. H. Whang, C. T. Liu, D. P. Pope and J. O. Stiegler, TMS, pp. 375-401, (1990)

/3/   *R. G. Rowe:* "The mechanical properties of titanium aluminides near Ti-25Al-25Nb", in Microstructure/Property Relationship in Titanium Aluminides and Alloys, eds. Y.-W. Kim and R. R. Boyer, TMS, pp. 387 - 398, (1991)

/4/   *P. R. Smith, J. A. Graves and C. G. Rhodes:* " Comparison of orthorhombic and alpha-two titanium aluminides as matrices for continuous SiC-reinforced composites ", Met. Trans., A., vol 25A, pp. 1267-1283, (1994)

/5/   *R. G. Rowe:* " Ti2AlNb-based alloys outperform conventional titanium aluminides ", Adv. Mater. Processes, vol 141, pp. 33-35, (1992)

/6/   *R. G. Rowe:* "The mechanical properties of ternary and quartenary Ti2NbAl-based titanium aluminide alloys", in Titanium '92 Science and Technology, eds. F. H. Froes and I. Caplan, TMS, pp. 343 - 350, (1993)

/7/ R. G. Rowe: "Tensile and creep behavior of ordered orthorhombic Ti2AlNb-based alloys", Mat. Res. Soc. Symp. Proc., 213, pp. 703 - 708, (1991)

/8/ J. A. Graves, P. R. Smith and C. G. Rhodes: " Evaluation of a Ti-22Al-23Nb orthorhombic alloy for use as the matrix in a high temperature Ti-based composite", in Intermetallic Matrix Composites II, eds. D. Miracle et al. , MRS, pp. 31 - 42, (1992)

/9/ G. A. Hartman and D. J. Buchanan: " Methodologies for thermal and mechanical testing of TMC material ", Characterization of fiber reinforced titanium matrix composites, AGARD Report 796, pp. 12-1 - 12-9, (1994)

/10/ G. A. Hartman and S. M. Russ: " Techniques for mechanical and thermal testing of Ti3Al/SCS-6 metal matrix composites ", Metal Matrix Composites: Testing, Analysis and Failure Modes, ASTM STP 1032, edited by W. J. Johnson, pp. 43-53, (1989)

/11/ C. G. Rhodes, J. A. Graves, P.R. Smith and M. R. James: "Characterization of orthorhombic titanium aluminide alloys", Structural Intermetallics, eds. R. Darbolia, J. J. Lewandowski, C. T. Lin, P. L. Martin, D. B. Miracle and M. V. Nathal, TMS, pp. 45 - 52, (1993)
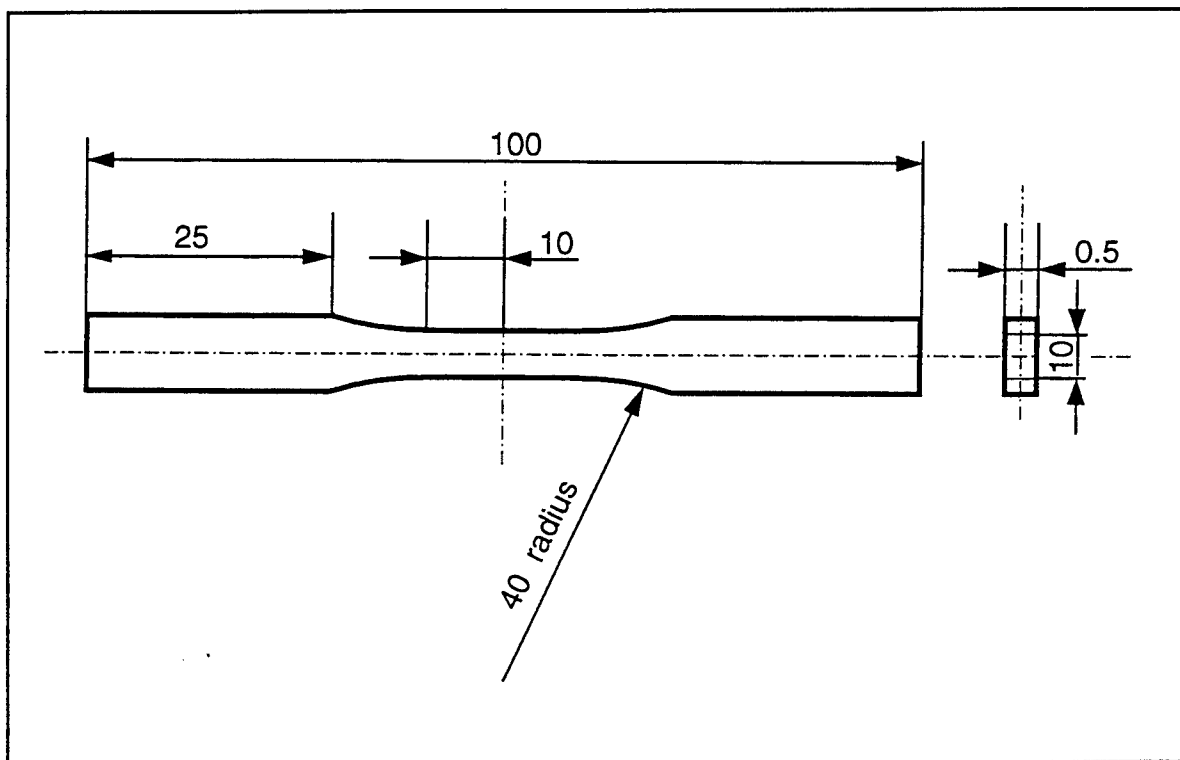
Figure 1: Schematic drawing of the fatigue specimen
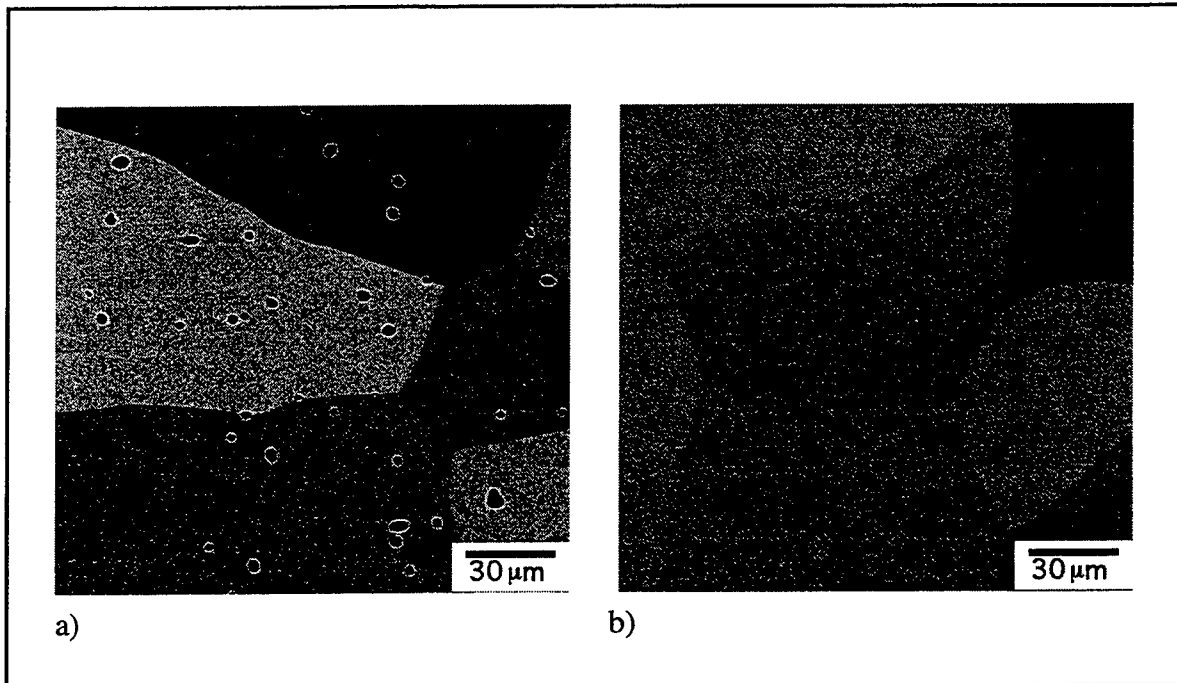
Figure 2: SEM micrograph of specimen, heat treated at a) 1060°C and b) 1080°C, each for 2 hours followed by an oil quench
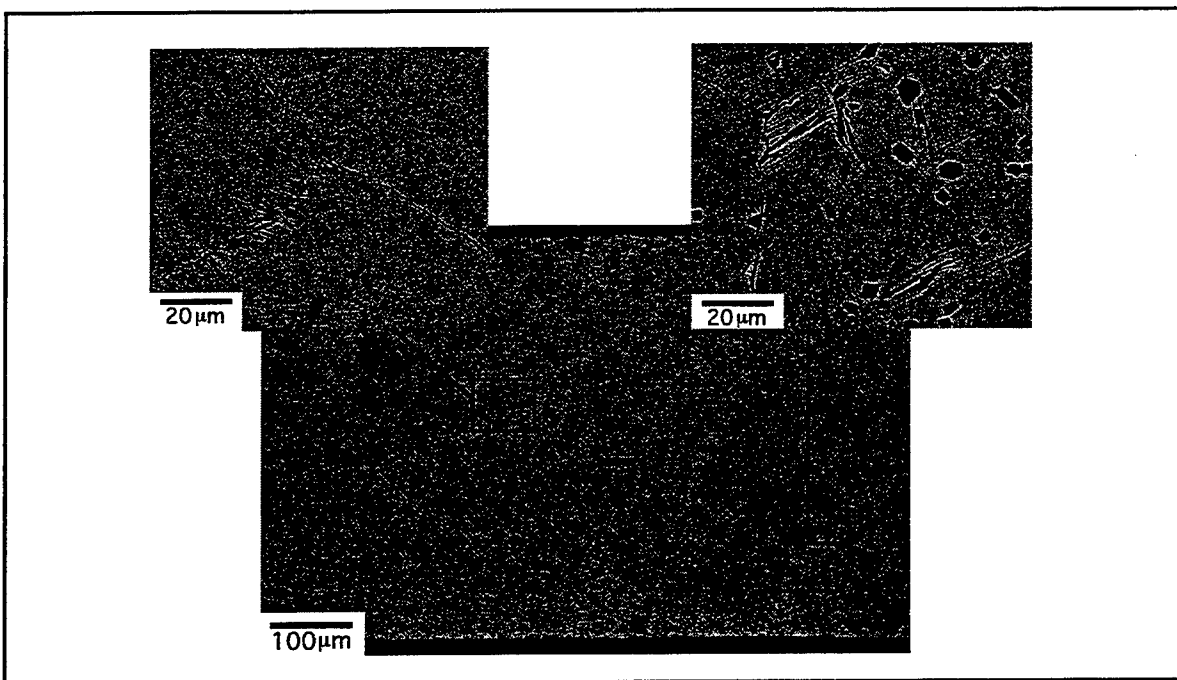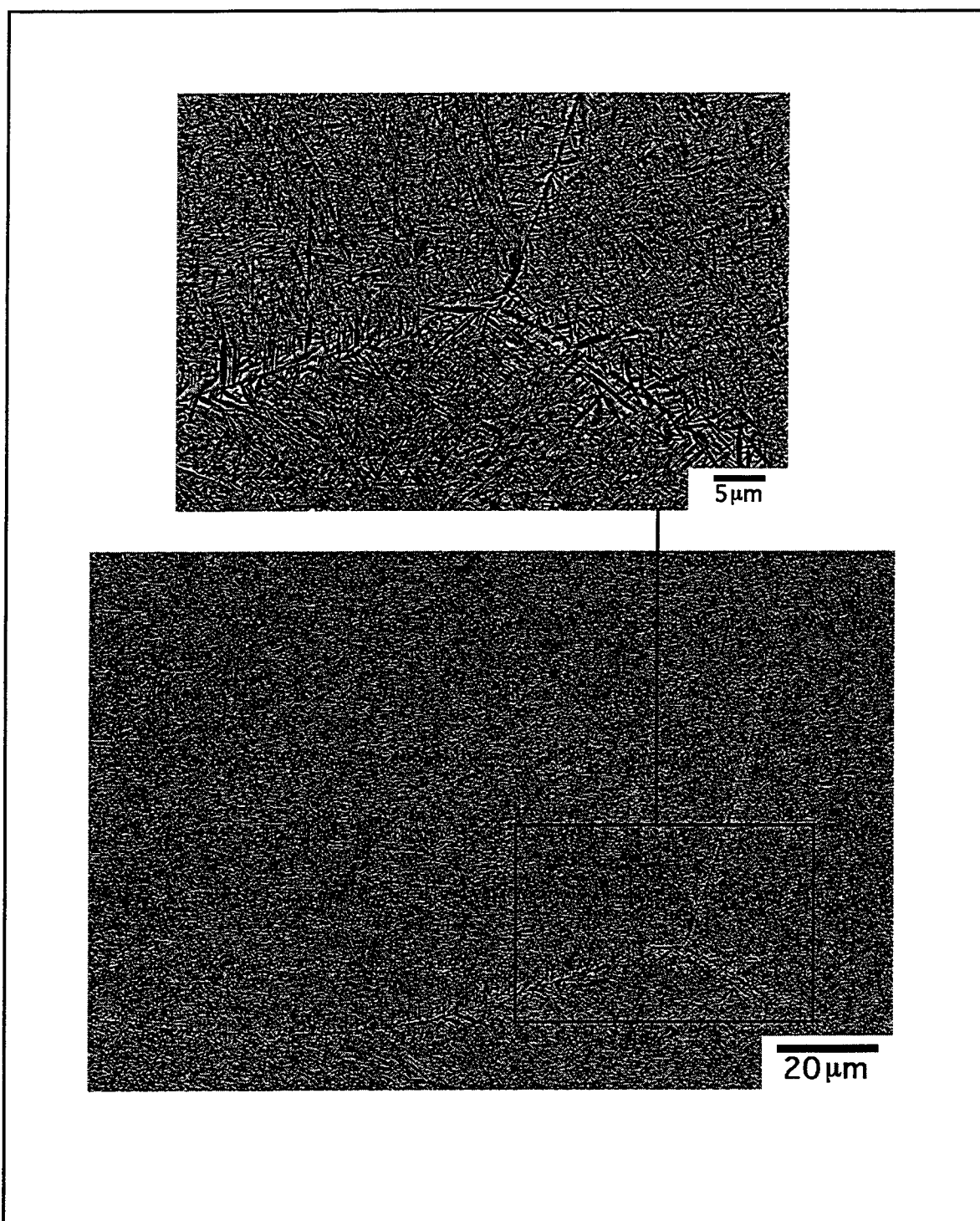

Figure 3: SEM micrograph of the specimen heat treated at 1075°C for 2 hours with a subsequent cooling rate of 28°C per min down to 815°C, where it was held for 8 hours and then furnace cooled
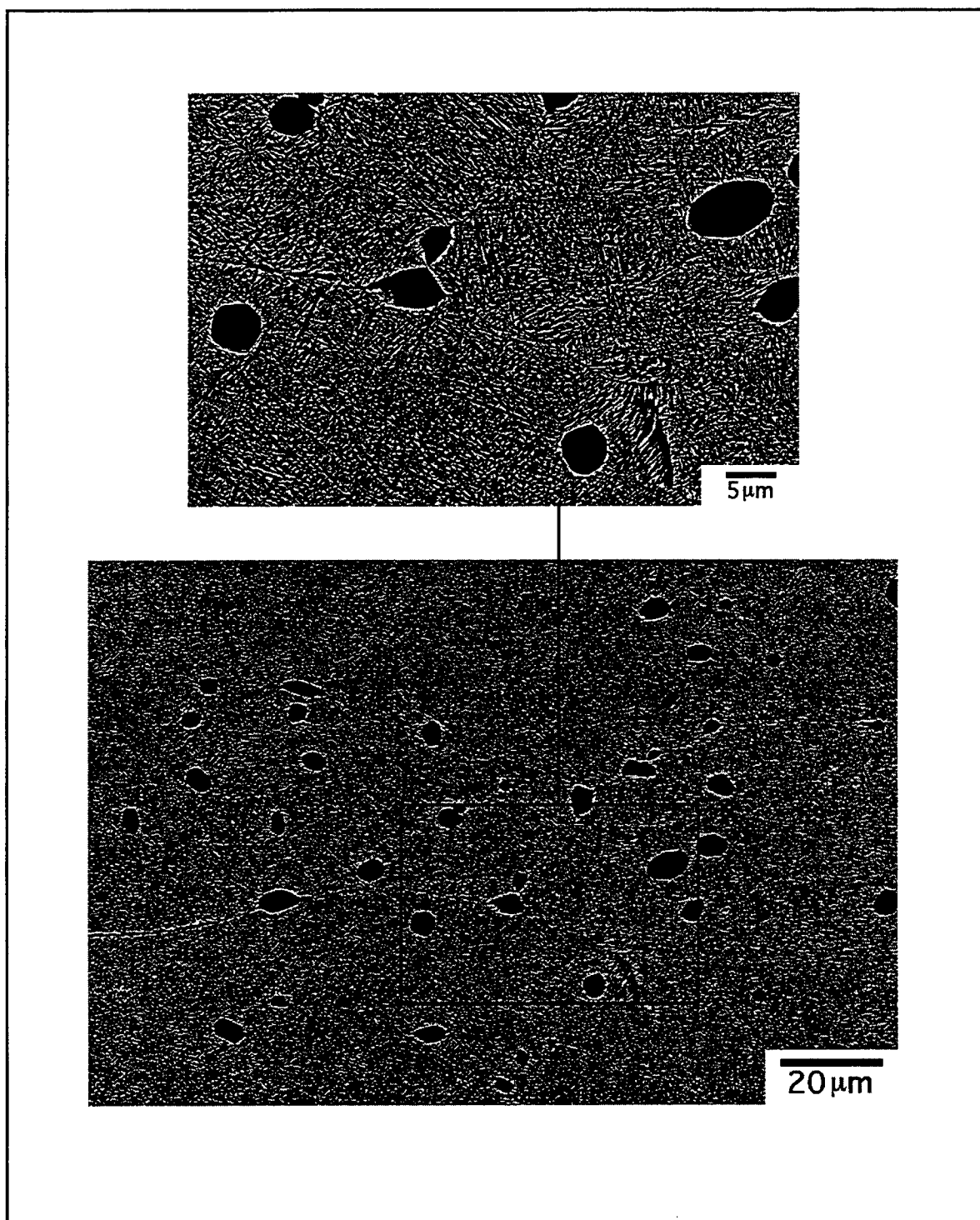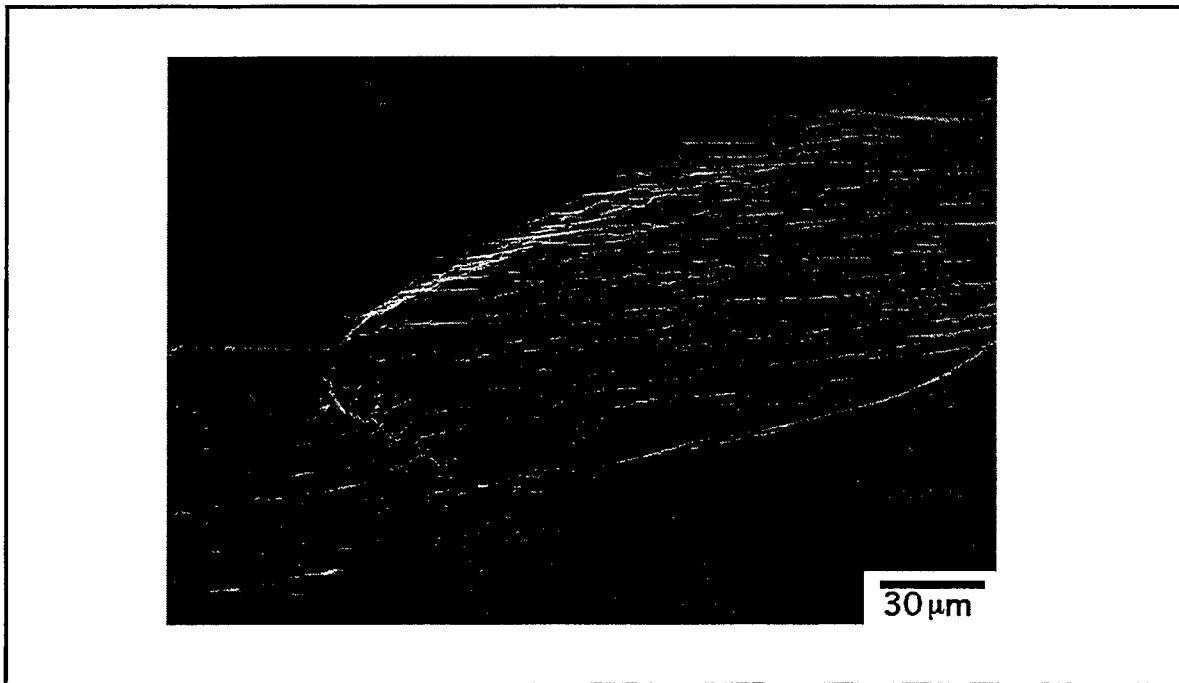
Figure 4: SEM micrograph of the specimen heat treated at 1055°C for 2 hours with a subsequent cooling rate of 50°C per min down to 815°C, where it was held for 8 hours and then furnace cooled, (microstructure A)

Figure 5: SEM micrograph of the specimen heat treated at 1095°C for 2 hours with a subsequent cooling rate of 50°C per min down to 815°C, where it was held for 8 hours and then furnace cooled, (microstructure B)

Figure 6: SEM micrograph of surface cracking in the specimen's gage section occurring during tensile testing at 540°C
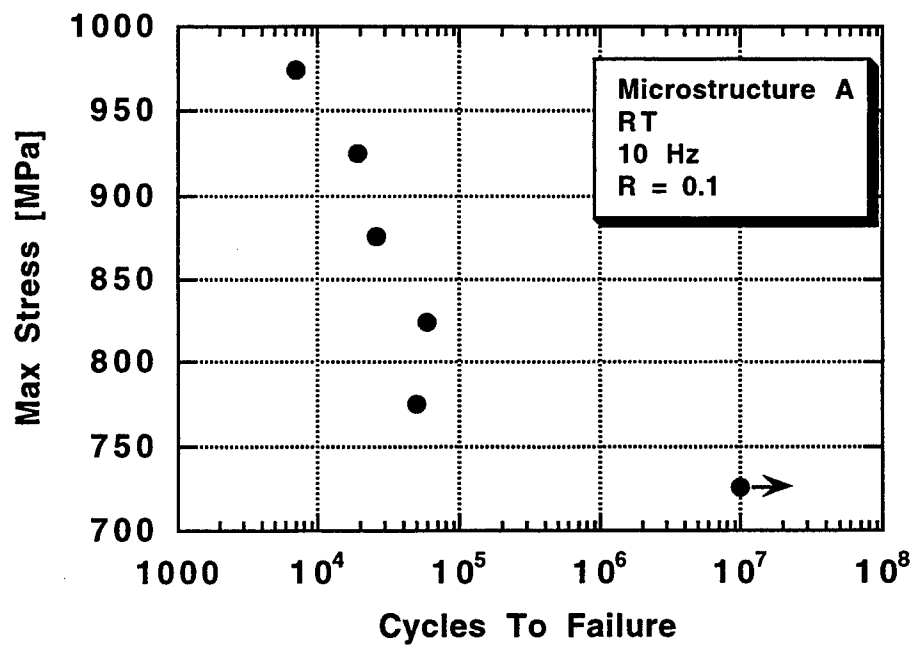
Figure 7: SN-curve of Ti-22Al-234Nb, microstructure A
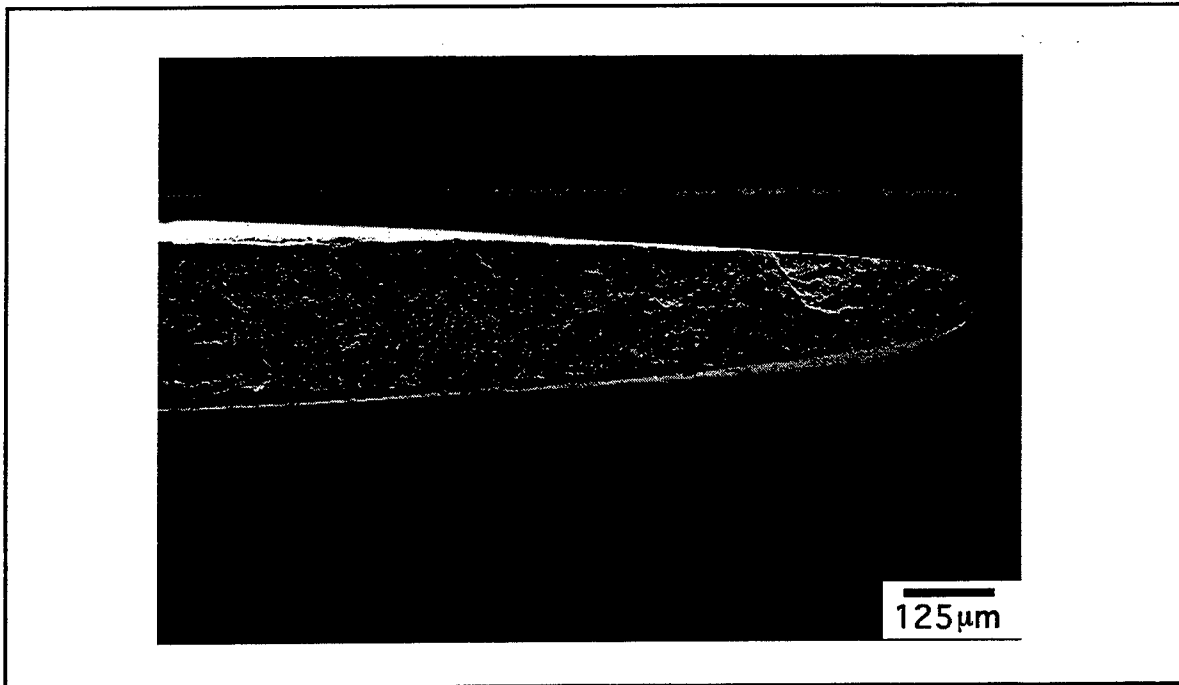
Figure 8: SEM micrograph of the fatigue specimen's fracture surface tested at 725 MPa at room temperature
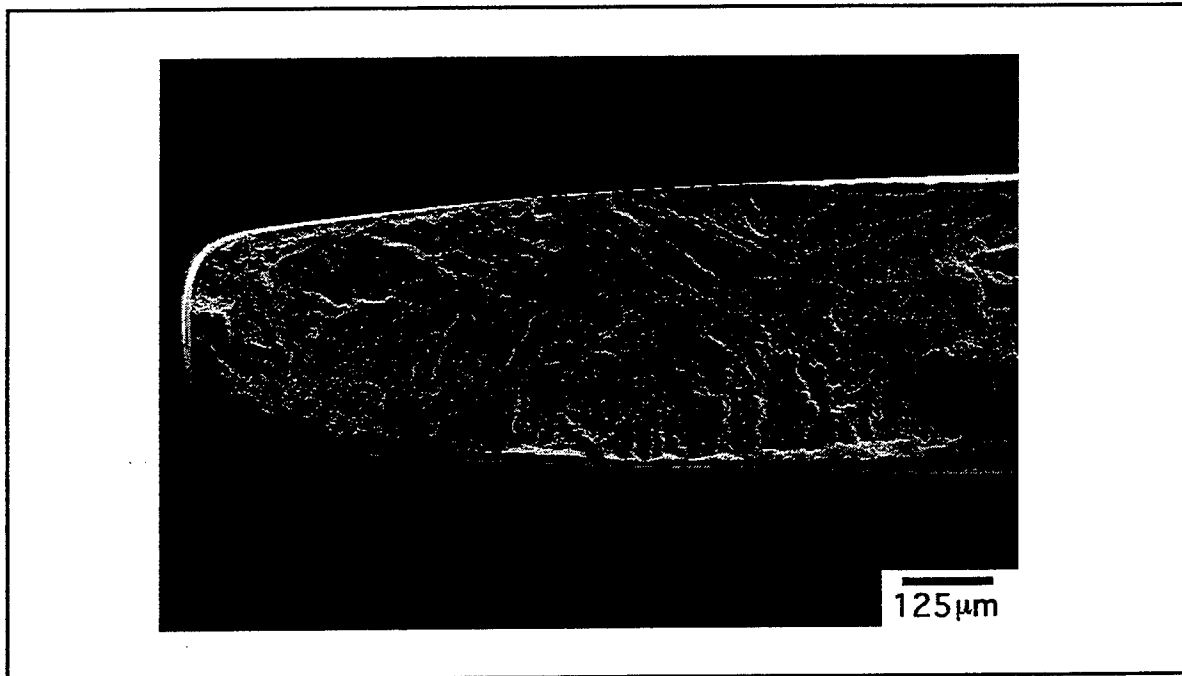
Figure 9: SEM micrograph of the fatigue specimen's fracture surface tested at 875 MPa at room temperature

Giovanni Luvera's report was not available at the time of publication.

CHARACTERIZATION OF SEMICONDUCTOR
JUNCTION IGNITOR DEVICE

Alfred L. Malone
Research Assistant
Department of Electrical Engineering

Auburn University
200 Broun Hall
Auburn, Al 36849

CHARACTERIZATION OF SEMICONDUCTOR JUNCTION IGNITOR DEVICE

Alfred L. Malone
Research Assistant
Department of Electrical Engineering
Auburn University

Abstract

Semiconductor junction ignitors (SJI) were characterized by analyzing the current and voltage measurements of the devices during excitation by high currents. The resistance, power, specific action and energy of the device as a function of time were determined. Calculations were made to determine the theoretical burst time of the device, which aided in determining the actual burst time. These are the preliminary characterizations of the SJI device that will be followed up by a more detailed study in the near future.

CHARACTERIZATION OF SEMICONDUCTOR JUNCTION IGNITOR DEVICE

Alfred L. Malone

Introduction

Semiconductor junction ignitors (SJI's) are monolithic electro-explosive devices that are capable of igniting pyrotechnic compounds using a capacitor discharge unit (CDU). These ignitors consist of pn junctions (diodes) electrically connected in series. The breakdown voltage of the diodes can be easily varied from a few tenths of a volt to several hundred. This electro-explosive device has been used to ignite various pyrotechnic compounds, such as boron potassium nitride (BKNO$_3$) and zirconium potassium perchlorate (ZPP).

Typically the ignitor is functioned using current generated by a capacitor discharge unit. As the capacitor discharges, the potential across the SJI exceeds the breakdown voltage of the pn junction and the device conducts current. As the current begins to flow, the junction region undergoes initially Peltier and then ohmic heating. Once the temperature of the Si exceeds the evaporation point, the junction vaporizes producing an ionized plasma that ignites the pyrotechnic compound. In order to better understand the functioning of the device, the plasma characteristics of the SJI must be determined. This is accomplished by examining some of the electronic characteristics, such as the burst time, burst current, specific action at burst, power dissipated at burst, energy required to activate the device, the resistance and the power stored in the device

during firing. The electronic characteristics were acquired by firing several devices at different voltages and monitoring the current, voltage and photonic discharge of the SJI's.

Methodology

The devices were fired using a high voltage fireset with a 0.2 μF capacitor. Several shots were taken at 750V, 1kV, 1.5kV and 2kV. These voltages were arbitrarily selected to simulate a range of possible real life voltages. This was done to make sure that the data was complete and reproducible.

After the device was mounted on the fireset, a high voltage pulse was produced to fire the SJI. The pulse was generated by a pulse generator and a high voltage source. When the pulse generator was triggered, a high voltage potential was placed across the SJI device thus allowing current to flow and functioning the device. The current was measured using a current viewing resistor. The voltage was acquired using a Techtronix 5200 high voltage probe. A photo diode was used to determine the intensity of the photonic emissions. All of the data was acquired by a Techtronix 648 digital oscilloscope. Once the data was acquired, it was saved using LabView, a graphical software package used in instrumentation that allowed the data from the oscilloscope to be processed. The data was then imported into a separate software package called VuPoint to be analyzed.

Results

The actual SJI device that was tested consisted of two junction regions connected in series. A picture was taken of the device before and after firing, shown in Figure 1a and b. In the first photo the narrow region between the two small blocks is the junction and the land area is covered with conductive epoxy. The second photo of the junction region shows the aftermath of the explosive plasma arc. This showed that the firing of the device not only destroys the junction but part of the silicon oxide and silicon substrate also. Note the land region also showed signs of plasma arcing. A layout of the SJI device is shown in Figure 2. The actual current was found by using a current viewing resistor (CVR), where the voltage measured was divided by the resistance of the CVR, $0.004837 \ \Omega$. The current trace shown in Figure 3 the amount current flowing through the device as a function of time. The current trace increased over approximately the first 250 ns and decreased over the next 250 ns. This supported the initial hypothesis that the current should increase pass burst because even though the junction has blown at this time, the current still has a path to flow through due to the plasma plume that is present. The decrease in current during the next 250 ns was due to the decay of the plasma plume, the plasma arc resistance changing and the RLC time constant of the circuit, which damps the current waveform.

Figure 4 illustrates the voltage across the SJI device as a function of time during firing. The voltage rapidly increases for the first 150 ns and then starts to decay. During the first 150 nanoseconds, there were four spikes in the voltage at roughly 26 ns, 50 ns,

100 ns, and 150 ns. The spikes are probably due to the burst of the junction, the breakdown of the air and the breakdown of the plasma plume.

The resistance curve depicts the resistance across the junction as a function of time. The resistance shown in Figure 5 was found by dividing the voltage by the current. The resistance decreases as the voltage is applied to the device. As the device explodes the resistance decreases. As the plasma plume decayed the resistance increased.

A calculation was made to determine the amount of energy to evaporate the junction. This calculation gave a good approximation of the burst energy (energy required to explode the SJI device). The equation used to determine the energy is as follows:

$$E = mC\Delta T_{melting} + mL_f + mC\Delta T_{boiling} + mL_v$$

where : mass, $m = 1.817E{-}8$ g

specific heat of aluminum, $C = 0.9$ J/gK

melting point of aluminum $= 933.5$ K

boiling point of aluminum $= 2767$ K

latent heat of fusion, $L_f = 397$ J/g

latent heat of vaporization, $L_v = 1.08E{-}4$ J/g

The evaluation of this equation gave the estimated burst energy as 0.0476 mJ. Knowing this, the burst time was easily extracted from the graph of the energy as being approximately 39 ns, corresponds to the first spike in the voltage curve. A graph of the energy versus time is shown in Figure 6.

The specific action is shown in Figure 7. The specific action is proportional to the power absorbed by the device. The specific action was used to determine the burst current density. For the SJI devices, the burst current density was approximately 31.33 kA^2/s.

Conclusions

The data for the semiconductor junction ignitor devices was analyzed. From the 1000 V shot data, a burst time of 39 ns was found. Knowing the burst time, the voltage at burst was found to be 67.51 V and the burst current was found to be 62.65 A. The energy required to blow the device was calculated to be 0.0476 mJ. These are just a few of the characterizations that will be found. In the future more detailed characterizations will be determined. A time history of the resistance could be determined and used in a PSPICE model of the device's electrical properties. Also a correlation can be made with streak camera photographs to better determine the plasma properties.
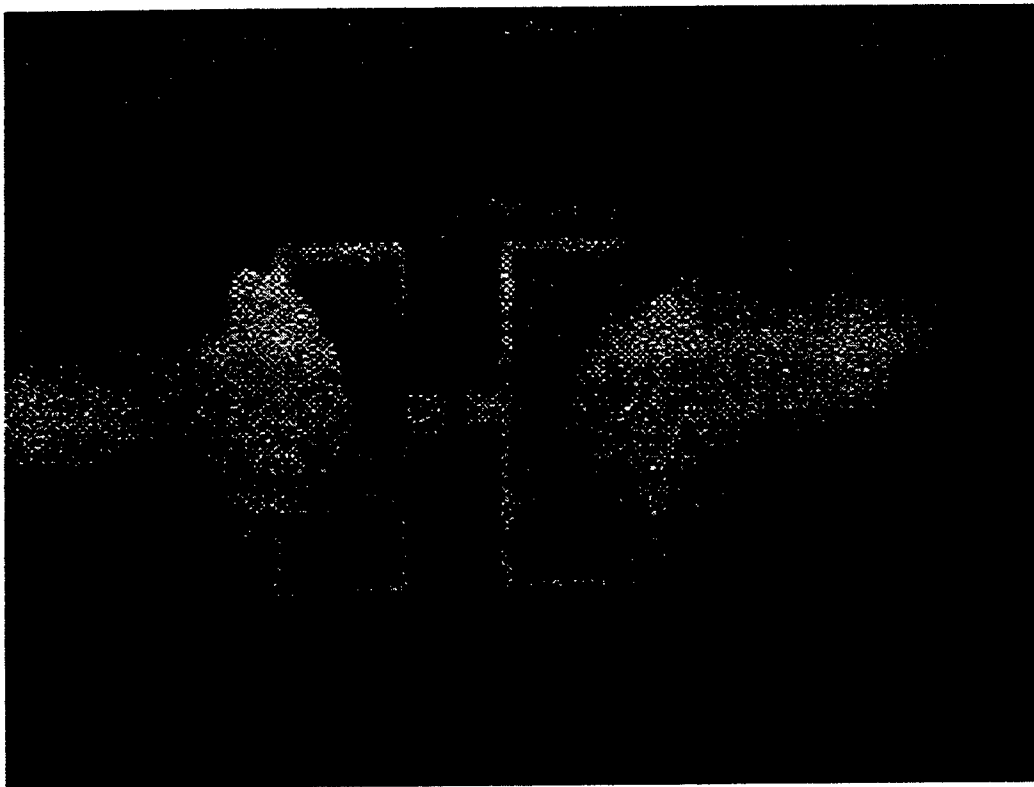
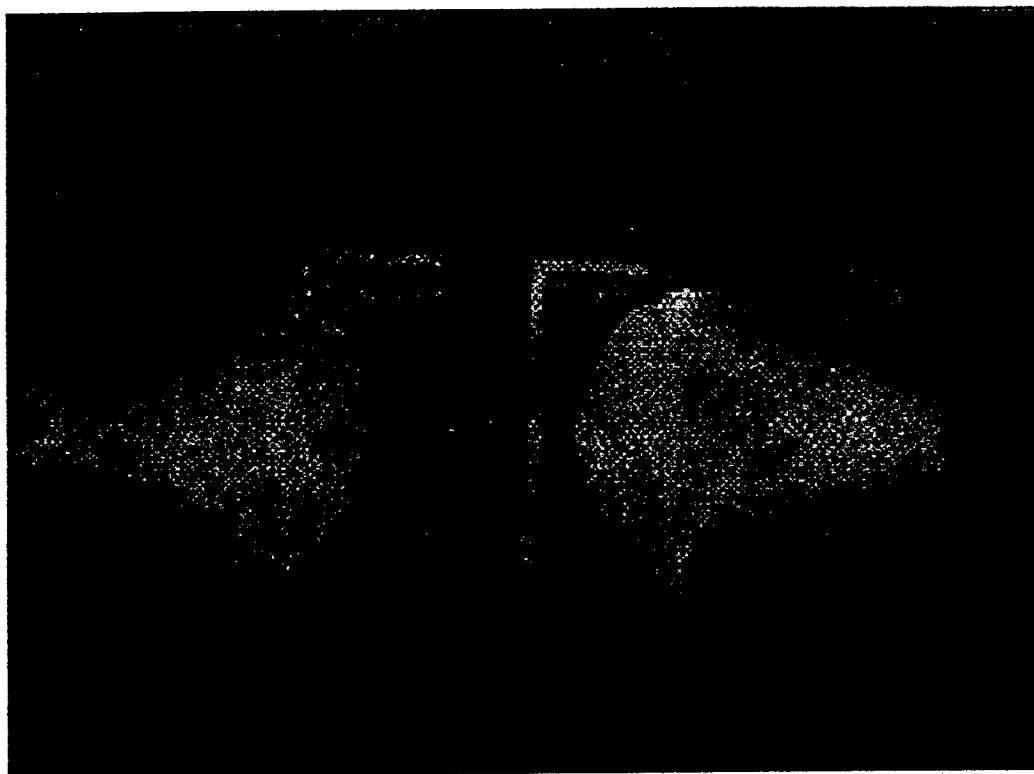Figure 1a : Semiconductor Junction Ignitor (Before Test)

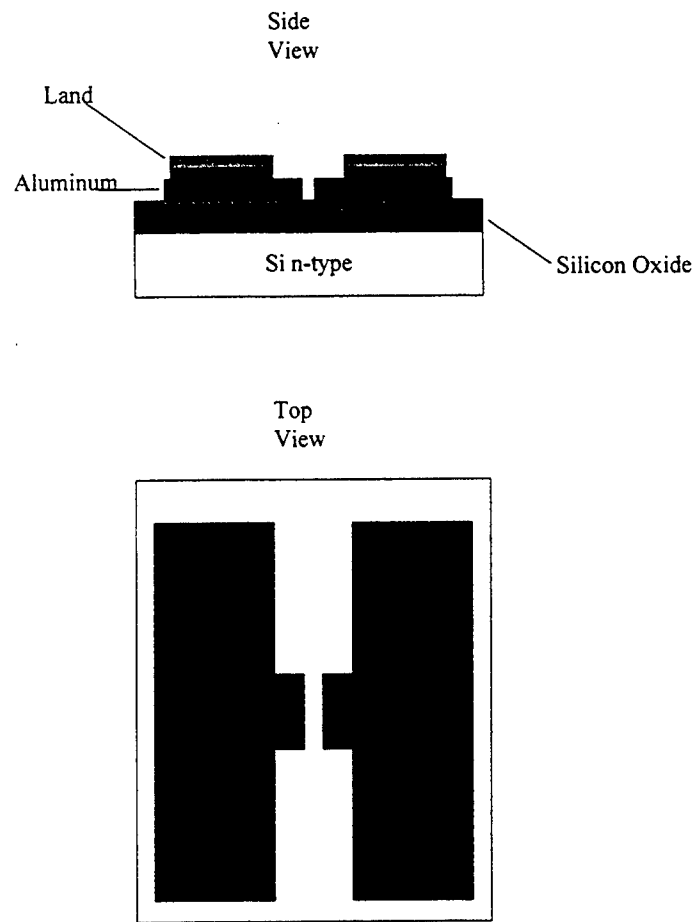Figure 1b : Semiconductor Junction Ignitor (After Test)

Side
View

Land

Aluminum

Si n-type

Silicon Oxide

Top
View

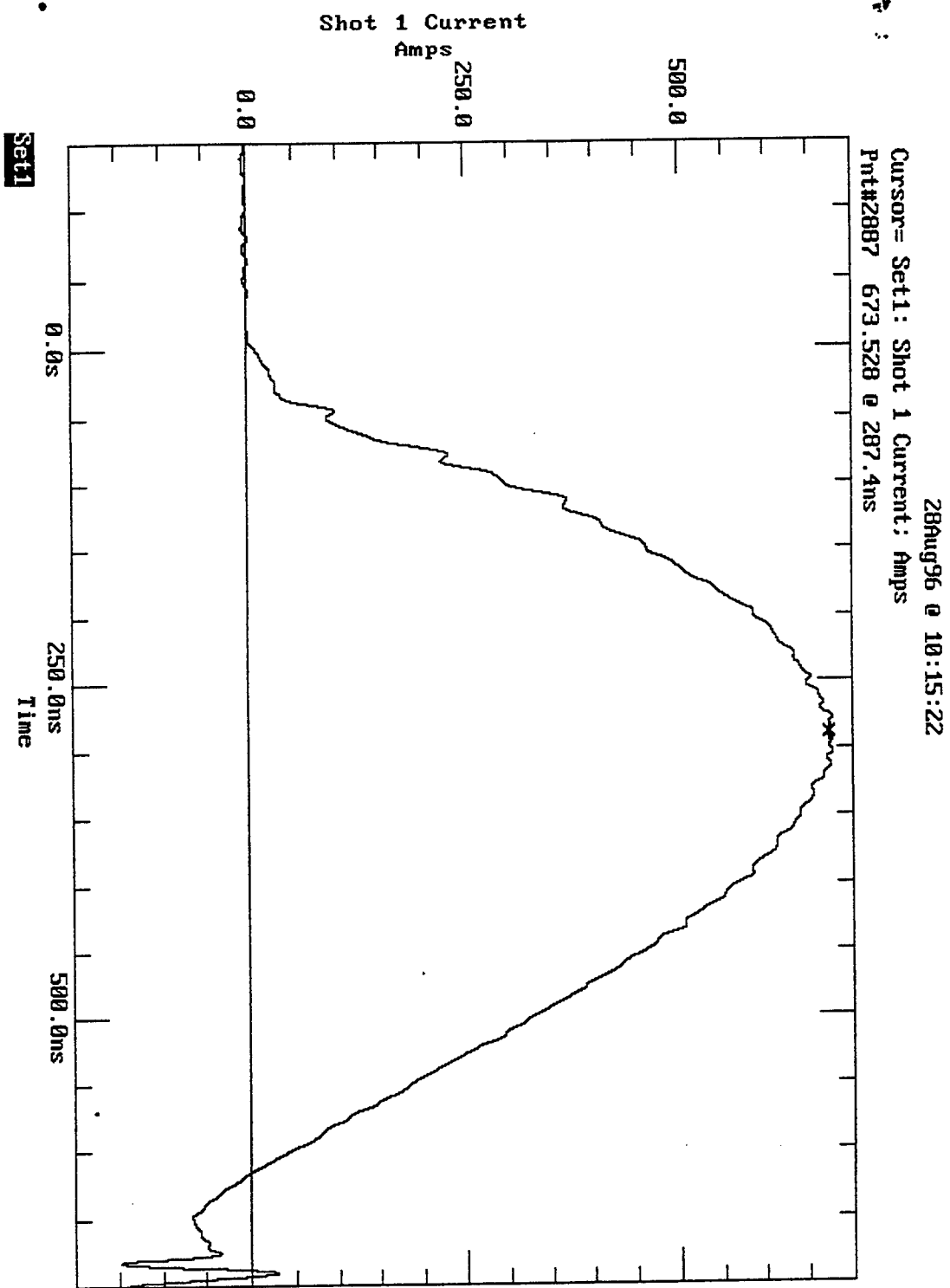Figure 2 : Physical Layout of the Semiconductor
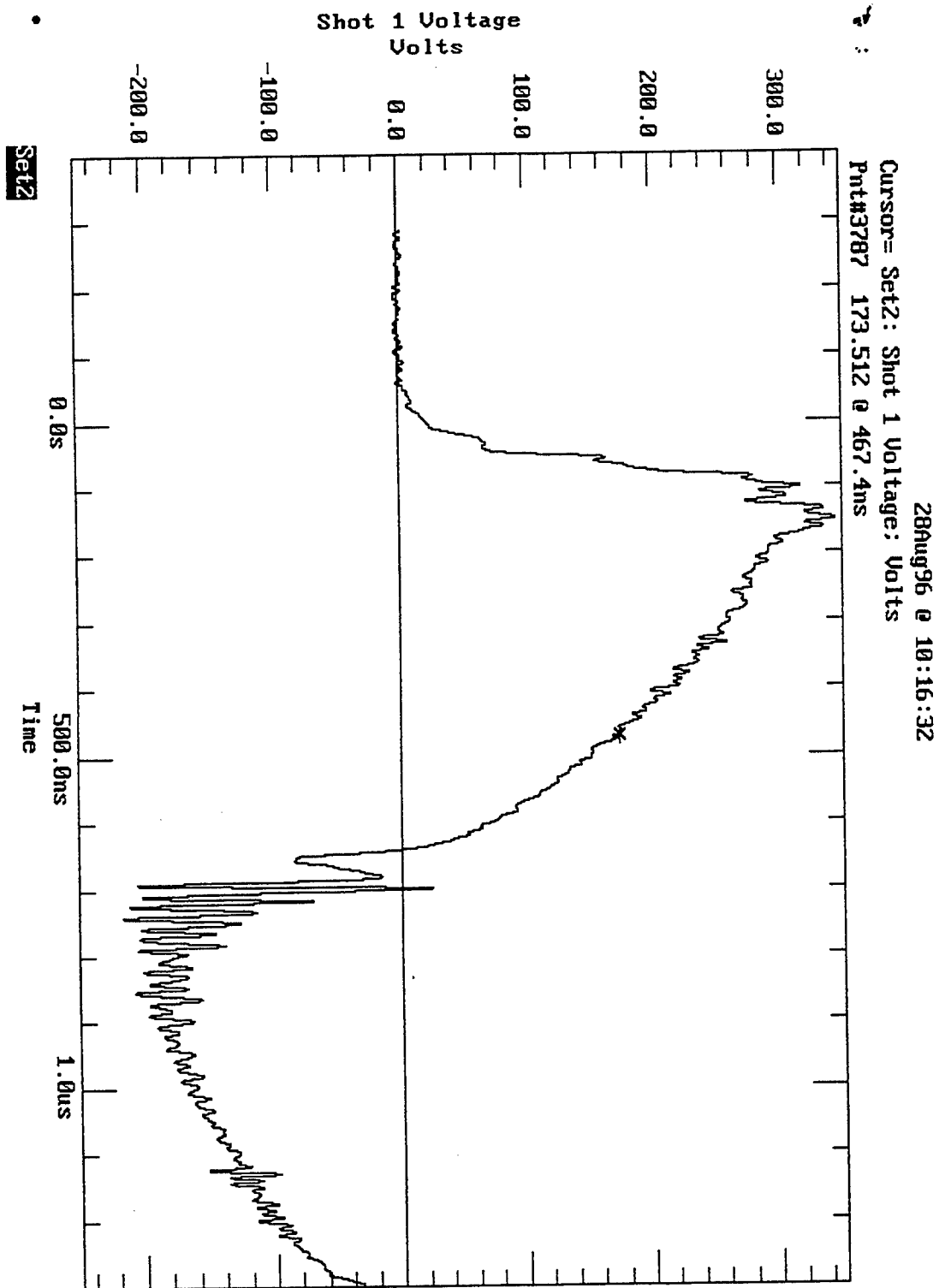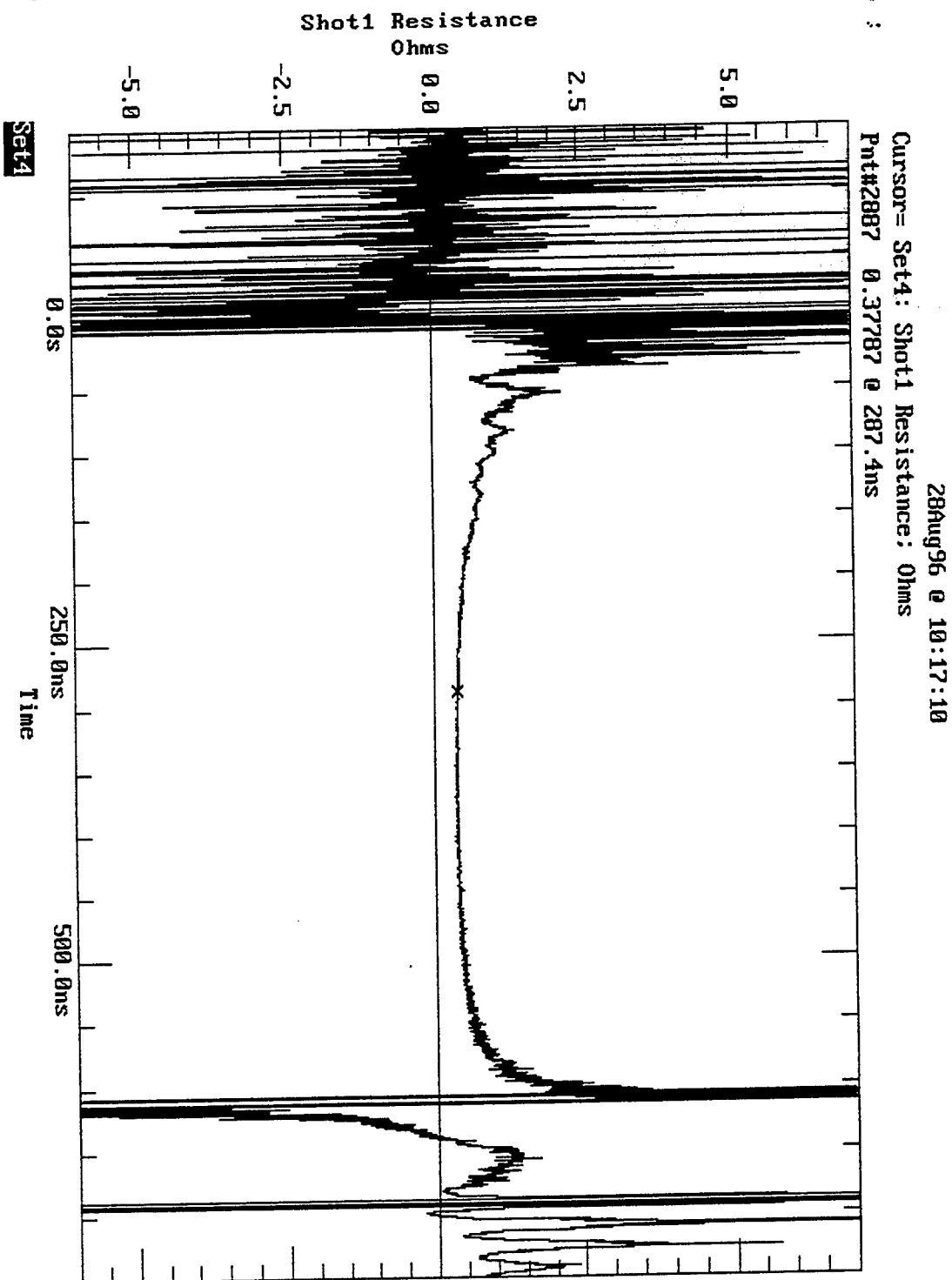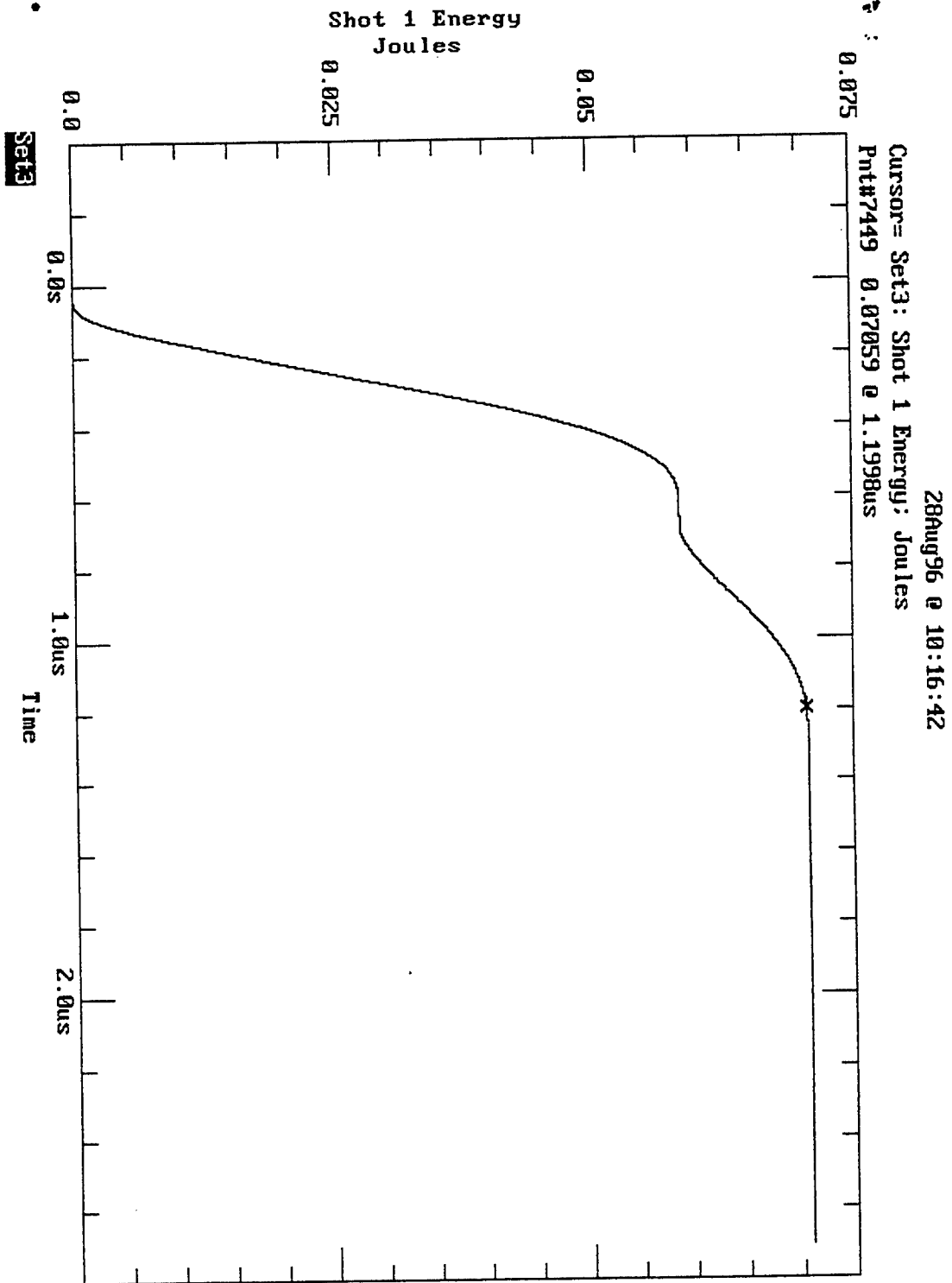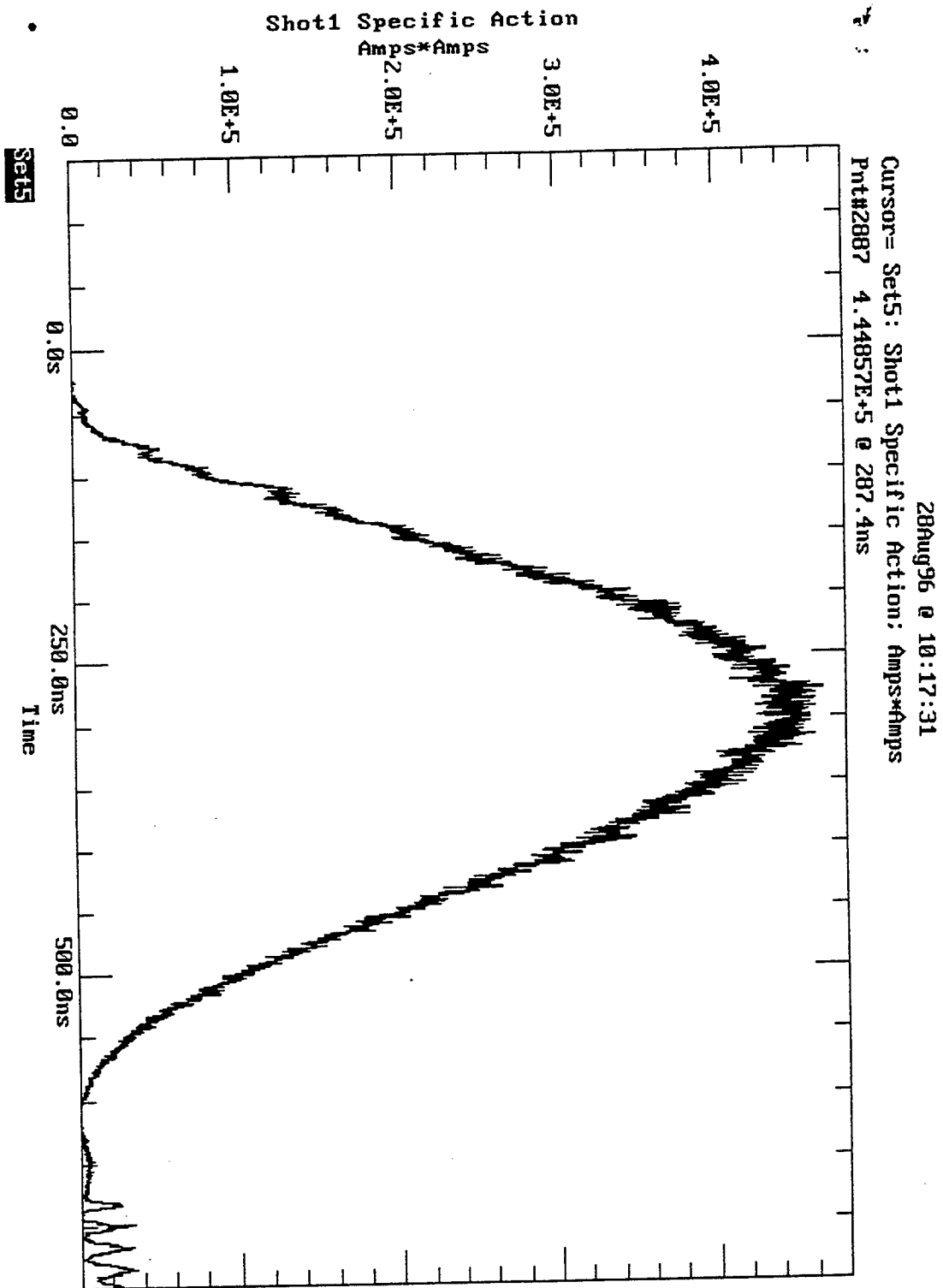Junction Ignitor Device

Figure 3: Current

Figure 4: Voltage

Figure 5: Resistance

Figure 6: Energy

Figure 7: Specific Action